Part IA — Vector Calculus

Based on lectures by B. Allanach

Notes taken by Dexter Chua

Lent 2015

These notes are not endorsed by the lecturers, and I have modified them (often significantly) after lectures. They are nowhere near accurate representations of what was actually lectured, and in particular, all errors are almost surely mine.

Curves in $\mathbb{R}^3$
Parameterised curves and arc length, tangents and normals to curves in $\mathbb{R}^3$, the radius of curvature. [1]

Integration in $\mathbb{R}^2$ and $\mathbb{R}^3$
Line integrals. Surface and volume integrals: definitions, examples using Cartesian, cylindrical and spherical coordinates; change of variables. [4]

Vector operators
Directional derivatives. The gradient of a real-valued function: definition; interpretation as normal to level surfaces; examples including the use of cylindrical, spherical *and general orthogonal curvilinear* coordinates.

Divergence, curl and ∇^2 in Cartesian coordinates, examples; formulae for these operators (statement only) in cylindrical, spherical *and general orthogonal curvilinear* coordinates. Solenoidal fields, irrotational fields and conservative fields; scalar potentials. Vector derivative identities. [5]

Integration theorems
Divergence theorem, Green's theorem, Stokes's theorem, Green's second theorem: statements; informal proofs; examples; application to fluid dynamics, and to electromagnetism including statement of Maxwell's equations. [5]

Laplace's equation
Laplace's equation in $\mathbb{R}^2$ and $\mathbb{R}^3$: uniqueness theorem and maximum principle. Solution of Poisson's equation by Gauss's method (for spherical and cylindrical symmetry) and as an integral. [4]

Cartesian tensors in $\mathbb{R}^3$
Tensor transformation laws, addition, multiplication, contraction, with emphasis on tensors of second rank. Isotropic second and third rank tensors. Symmetric and antisymmetric tensors. Revision of principal axes and diagonalization. Quotient theorem. Examples including inertia and conductivity. [5]

Contents

0 Introduction **4**

1 Derivatives and coordinates **5**
1.1 Derivative of functions 5
1.2 Inverse functions 9
1.3 Coordinate systems 10

2 Curves and Line **11**
2.1 Parametrised curves, lengths and arc length 11
2.2 Line integrals of vector fields 12
2.3 Gradients and Differentials 14
2.4 Work and potential energy 15

3 Integration in $\mathbb{R}^2$ and $\mathbb{R}^3$ **17**
3.1 Integrals over subsets of $\mathbb{R}^2$ 17
3.2 Change of variables for an integral in $\mathbb{R}^2$ 19
3.3 Generalization to $\mathbb{R}^3$ 21
3.4 Further generalizations 24

4 Surfaces and surface integrals **26**
4.1 Surfaces and Normal 26
4.2 Parametrized surfaces and area 27
4.3 Surface integral of vector fields 29
4.4 Change of variables in $\mathbb{R}^2$ and $\mathbb{R}^3$ revisited 31

5 Geometry of curves and surfaces **32**

6 Div, Grad, Curl and ∇ **35**
6.1 Div, Grad, Curl and ∇ 35
6.2 Second-order derivatives 37

7 Integral theorems **38**
7.1 Statement and examples 38
7.1.1 Green's theorem (in the plane) 38
7.1.2 Stokes' theorem 39
7.1.3 Divergence/Gauss theorem 40
7.2 Relating and proving integral theorems 41

8 Some applications of integral theorems **46**
8.1 Integral expressions for div and curl 46
8.2 Conservative fields and scalar products 47
8.3 Conservation laws 49

9 Orthogonal curvilinear coordinates **51**
9.1 Line, area and volume elements 51
9.2 Grad, Div and Curl 52

10 Gauss' Law and Poisson's equation **54**
10.1 Laws of gravitation 54
10.2 Laws of electrostatics 55
10.3 Poisson's Equation and Laplace's equation 57

11 Laplace's and Poisson's equations **61**
11.1 Uniqueness theorems 61
11.2 Laplace's equation and harmonic functions 62
11.2.1 The mean value property 62
11.2.2 The maximum (or minimum) principle 63
11.3 Integral solutions of Poisson's equations 64
11.3.1 Statement and informal derivation 64
11.3.2 Point sources and δ-functions* 65

12 Maxwell's equations **67**
12.1 Laws of electromagnetism 67
12.2 Static charges and steady currents 68
12.3 Electromagnetic waves 69

13 Tensors and tensor fields **70**
13.1 Definition 70
13.2 Tensor algebra 71
13.3 Symmetric and antisymmetric tensors 72
13.4 Tensors, multi-linear maps and the quotient rule 73
13.5 Tensor calculus 74

14 Tensors of rank 2 **77**
14.1 Decomposition of a second-rank tensor 77
14.2 The inertia tensor 78
14.3 Diagonalization of a symmetric second rank tensor 80

15 Invariant and isotropic tensors **81**
15.1 Definitions and classification results 81
15.2 Application to invariant integrals 82

0 Introduction

In the differential equations class, we learnt how to do calculus in one dimension. However, (apparently) the world has more than one dimension. We live in a 3 (or 4) dimensional world, and string theorists think that the world has more than 10 dimensions. It is thus important to know how to do calculus in many dimensions.

For example, the position of a particle in a three dimensional world can be given by a position vector $\mathbf{x}$. Then by definition, the velocity is given by $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x} = \dot{\mathbf{x}}$. This would require us to take the derivative of a vector.

This is not too difficult. We can just differentiate the vector componentwise. However, we can reverse the problem and get a more complicated one. We can assign a number to each point in (3D) space, and ask how this number changes as we move in space. For example, the function might tell us the temperature at each point in space, and we want to know how the temperature changes with position.

In the most general case, we will assign a vector to each point in space. For example, the electric field vector $\mathbf{E}(\mathbf{x})$ tells us the direction of the electric field at each point in space.

On the other side of the story, we also want to do integration in multiple dimensions. Apart from the obvious "integrating a vector", we might want to integrate over surfaces. For example, we can let $\mathbf{v}(\mathbf{x})$ be the velocity of some fluid at each point in space. Then to find the total fluid flow through a surface, we integrate $\mathbf{v}$ over the surface.

In this course, we are mostly going to learn about doing calculus in many dimensions. In the last few lectures, we are going to learn about Cartesian tensors, which is a generalization of vectors.

Note that throughout the course (and lecture notes), summation convention is implied unless otherwise stated.

1 Derivatives and coordinates

1.1 Derivative of functions

We used to define a derivative as the limit of a quotient and a function is differentiable if the derivative exists. However, this obviously cannot be generalized to vector-valued functions, since you cannot divide by vectors. So we want an alternative definition of differentiation, which can be easily generalized to vectors.

Recall, that if a function f is differentiable at x, then for a small perturbation δx, we have

$$\delta f \stackrel{\text{def}}{=} f(x+\delta x) - f(x) = f'(x)\delta x + o(\delta x),$$

which says that the resulting change in f is approximately proportional to δx (as opposed to $1/\delta x$ or something else). It can be easily shown that the converse is true — if f satisfies this relation, then f is differentiable.

This definition is more easily extended to vector functions. We say a function $\mathbf{F}$ is differentiable if, when x is perturbed by δx, then the resulting change is "something" times δx plus an $o(\delta x)$ error term. In the most general case, δx will be a vector and that "something" will be a matrix. Then that "something" will be what we call the derivative.

Vector functions $\mathbb{R} \to \mathbb{R}^n$

We start with the simple case of vector functions.

Definition (Vector function)**.** A *vector function* is a function $\mathbf{F} : \mathbb{R} \to \mathbb{R}^n$.

This takes in a number and returns a vector. For example, it can map a time to the velocity of a particle at that time.

Definition (Derivative of vector function)**.** A vector function $\mathbf{F}(x)$ is *differentiable* if

$$\delta\mathbf{F} \stackrel{\text{def}}{=} \mathbf{F}(x+\delta x) - \mathbf{F}(x) = \mathbf{F}'(x)\delta x + o(\delta x)$$

for some $\mathbf{F}'(x)$. $\mathbf{F}'(x)$ is called the *derivative* of $\mathbf{F}(x)$.

We don't have anything new and special here, since we might as well have defined $\mathbf{F}'(x)$ as

$$\mathbf{F}' = \frac{\mathrm{d}\mathbf{F}}{\mathrm{d}x} = \lim_{\delta x \to 0} \frac{1}{\delta x}[\mathbf{F}(x+\delta x) - \mathbf{F}(x)],$$

which is easily shown to be equivalent to the above definition.

Using differential notation, the differentiability condition can be written as

$$\mathrm{d}\mathbf{F} = \mathbf{F}'(x)\,\mathrm{d}x.$$

Given a basis $\mathbf{e}_i$ that is independent of x, vector differentiation is performed componentwise, i.e.

Proposition.

$$\mathbf{F}'(x) = F_i'(x)\mathbf{e}_i.$$

Leibnitz identities hold for the products of scalar and vector functions.

Proposition.

$$\frac{\mathrm{d}}{\mathrm{d}t}(f\mathbf{g}) = \frac{\mathrm{d}f}{\mathrm{d}t}\mathbf{g} + f\frac{\mathrm{d}\mathbf{g}}{\mathrm{d}t}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{g}\cdot\mathbf{h}) = \frac{\mathrm{d}\mathbf{g}}{\mathrm{d}t}\cdot\mathbf{h} + \mathbf{g}\cdot\frac{\mathrm{d}\mathbf{h}}{\mathrm{d}t}$$
$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{g}\times\mathbf{h}) = \frac{\mathrm{d}\mathbf{g}}{\mathrm{d}t}\times\mathbf{h} + \mathbf{g}\times\frac{\mathrm{d}\mathbf{h}}{\mathrm{d}t}$$

Note that the order of multiplication must be retained in the case of the cross product.

Example. Consider a particle with mass m. It has position $\mathbf{r}(t)$, velocity $\dot{\mathbf{r}}(t)$ and acceleration $\ddot{\mathbf{r}}$. Its momentum is $\mathbf{p} = m\dot{\mathbf{r}}(t)$.

Note that derivatives with respect to t are usually denoted by dots instead of dashes.

If $\mathbf{F}(\mathbf{r})$ is the force on a particle, then Newton's second law states that

$$\dot{\mathbf{p}} = m\ddot{\mathbf{r}} = \mathbf{F}.$$

We can define the angular momentum about the origin to be

$$\mathbf{L} = \mathbf{r}\times\mathbf{p} = m\mathbf{r}\times\dot{\mathbf{r}}.$$

If we want to know how the angular momentum changes over time, then

$$\dot{\mathbf{L}} = m\dot{\mathbf{r}}\times\dot{\mathbf{r}} + m\mathbf{r}\times\ddot{\mathbf{r}} = m\mathbf{r}\times\ddot{\mathbf{r}} = \mathbf{r}\times\mathbf{F}.$$

which is the *torque* of $\mathbf{F}$ about the origin.

Scalar functions $\mathbb{R}^n \to \mathbb{R}$

We can also define derivatives for a different kind of function:

Definition. A *scalar function* is a function $f : \mathbb{R}^n \to \mathbb{R}$.

A scalar function takes in a position and gives you a number, e.g. the potential energy of a particle at different positions.

Before we define the derivative of a scalar function, we have to first define what it means to take a limit of a vector.

Definition (Limit of vector)**.** The *limit of vectors* is defined using the norm. So $\mathbf{v}\to\mathbf{c}$ iff $|\mathbf{v}-\mathbf{c}|\to 0$. Similarly, $f(\mathbf{r}) = o(\mathbf{r})$ means $\frac{|f(\mathbf{r})|}{|\mathbf{r}|}\to 0$ as $\mathbf{r}\to\mathbf{0}$.

Definition (Gradient of scalar function)**.** A scalar function $f(\mathbf{r})$ is *differentiable* at $\mathbf{r}$ if

$$\delta f \stackrel{\text{def}}{=} f(\mathbf{r}+\delta\mathbf{r}) - f(\mathbf{r}) = (\nabla f)\cdot\delta\mathbf{r} + o(\delta\mathbf{r})$$

for some vector ∇f, the *gradient* of f at $\mathbf{r}$.

Here we have a fancy name "gradient" for the derivative. But we will soon give up on finding fancy names and just call everything the "derivative"!

Note also that here we genuinely need the new notion of derivative, since "dividing by $\delta\mathbf{r}$" makes no sense at all!

The above definition considers the case where $\delta\mathbf{r}$ comes in all directions. What if we only care about the case where $\delta\mathbf{r}$ is in some particular direction $\mathbf{n}$? For example, maybe f is the potential of a particle that is confined to move in one straight line only.

Then taking $\delta\mathbf{r} = h\mathbf{n}$, with $\mathbf{n}$ a unit vector,

$$f(\mathbf{r} + h\mathbf{n}) - f(\mathbf{r}) = \nabla f \cdot (h\mathbf{n}) + o(h) = h(\nabla f \cdot \mathbf{n}) + o(h),$$

which gives

Definition (Directional derivative)**.** The *directional derivative* of f along $\mathbf{n}$ is

$$\mathbf{n} \cdot \nabla f = \lim_{h\to 0} \frac{1}{h}[f(\mathbf{r} + h\mathbf{n}) - f(\mathbf{r})],$$

It refers to how fast f changes when we move in the direction of $\mathbf{n}$.

Using this expression, the directional derivative is maximized when $\mathbf{n}$ is in the same direction as ∇f (then $\mathbf{n} \cdot \nabla f = |\nabla f|$). So ∇f points in the direction of greatest slope.

How do we evaluate ∇f? Suppose we have an orthonormal basis $\mathbf{e}_i$. Setting $\mathbf{n} = \mathbf{e}_i$ in the above equation, we obtain

$$\mathbf{e}_i \cdot \nabla f = \lim_{h\to 0} \frac{1}{h}[f(\mathbf{r} + h\mathbf{e}_i) - f(\mathbf{r})] = \frac{\partial f}{\partial x_i}.$$

Hence

Theorem. The gradient is

$$\nabla f = \frac{\partial f}{\partial x_i}\mathbf{e}_i$$

Hence we can write the condition of differentiability as

$$\delta f = \frac{\partial f}{\partial x_i}\delta x_i + o(\delta\mathbf{x}).$$

In differential notation, we write

$$\mathrm{d}f = \nabla f \cdot \mathrm{d}\mathbf{r} = \frac{\partial f}{\partial x_i}\mathrm{d}x_i,$$

which is the chain rule for partial derivatives.

Example. Take $f(x, y, z) = x + e^{xy}\sin z$. Then

$$\begin{aligned}\nabla f &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)\\ &= (1 + ye^{xy}\sin z, xe^{xy}\sin z, e^{xy}\cos z)\end{aligned}$$

At $(x, y, z) = (0, 1, 0)$, $\nabla f = (1, 0, 1)$. So f increases/decreases most rapidly for $\mathbf{n} = \pm\frac{1}{\sqrt{2}}(1, 0, 1)$ with a rate of change of $\pm\sqrt{2}$. There is no change in f if $\mathbf{n}$ is perpendicular to $\pm\frac{1}{\sqrt{2}}(1, 0, 1)$.

Now suppose we have a scalar function $f(\mathbf{r})$ and we want to consider the rate of change along a path $\mathbf{r}(u)$. A change δu produces a change $\delta\mathbf{r} = \mathbf{r}'\delta u + o(\delta u)$, and

$$\delta f = \nabla f \cdot \delta\mathbf{r} + o(|\delta\mathbf{r}|) = \nabla f \cdot \mathbf{r}'(u)\delta u + o(\delta u).$$

This shows that f is differentiable as a function of u and

Theorem (Chain rule)**.** Given a function $f(\mathbf{r}(u))$,

$$\frac{\mathrm{d}f}{\mathrm{d}u} = \nabla f \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}u} = \frac{\partial f}{\partial x_i}\frac{\mathrm{d}x_i}{\mathrm{d}u}.$$

Note that if we drop the $\mathrm{d}u$, we simply get

$$\mathrm{d}f = \nabla f \cdot \mathrm{d}\mathbf{r} = \frac{\partial f}{\partial x_i}\mathrm{d}x_i,$$

which is what we've previously had.

Vector fields $\mathbb{R}^n \to \mathbb{R}^m$

We are now ready to tackle the general case, which are given the fancy name of *vector fields*.

Definition (Vector field)**.** A *vector field* is a function $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$.

Definition (Derivative of vector field)**.** A vector field $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable if

$$\delta\mathbf{F} \overset{def}{=} \mathbf{F}(\mathbf{x} + \delta\mathbf{x}) - \mathbf{F}(\mathbf{x}) = M\delta\mathbf{x} + o(\delta\mathbf{x})$$

for some $m \times n$ matrix M. M is the *derivative* of $\mathbf{F}$.

As promised, M does not have a fancy name.

Given an arbitrary function $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$ that maps $\mathbf{x} \mapsto \mathbf{y}$ and a choice of basis, we can write $\mathbf{F}$ as a set of m functions $y_j = F_j(\mathbf{x})$ such that $\mathbf{y} = (y_1, y_2, \cdots, y_m)$. Then

$$\mathrm{d}y_j = \frac{\partial F_j}{\partial x_i}\mathrm{d}x_i.$$

and we can write the derivative as

Theorem. The derivative of $\mathbf{F}$ is given by

$$M_{ji} = \frac{\partial y_j}{\partial x_i}.$$

Note that we could have used this as the definition of the derivative. However, the original definition is superior because it does not require a selection of coordinate system.

Definition. A function is *smooth* if it can be differentiated any number of times. This requires that all partial derivatives exist and are totally symmetric in i, j and k (i.e. the differential operator is commutative).

The functions we will consider will be smooth except where things obviously go wrong (e.g. $f(x) = 1/x$ at $x = 0$).

Theorem (Chain rule)**.** Suppose $g : \mathbb{R}^p \to \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^m$. Suppose that the coordinates of the vectors in $\mathbb{R}^p, \mathbb{R}^n$ and $\mathbb{R}^m$ are u_a, x_i and y_r respectively. By the chain rule,

$$\frac{\partial y_r}{\partial u_a} = \frac{\partial y_r}{\partial x_i}\frac{\partial x_i}{\partial u_a},$$

with summation implied. Writing in matrix form,

$$M(f \circ g)_{ra} = M(f)_{ri} M(g)_{ia}.$$

Alternatively, in operator form,

$$\frac{\partial}{\partial u_a} = \frac{\partial x_i}{\partial u_a}\frac{\partial}{\partial x_i}.$$

1.2 Inverse functions

Suppose $g, f : \mathbb{R}^n \to \mathbb{R}^n$ are inverse functions, i.e. $g \circ f = f \circ g = \mathrm{id}$. Suppose that $f(\mathbf{x}) = \mathbf{u}$ and $g(\mathbf{u}) = \mathbf{x}$.

Since the derivative of the identity function is the identity matrix (if you differentiate $\mathbf{x}$ wrt to $\mathbf{x}$, you get 1), we must have

$$M(f \circ g) = I.$$

Therefore we know that

$$M(g) = M(f)^{-1}.$$

We derive this result more formally by noting

$$\frac{\partial u_b}{\partial u_a} = \delta_{ab}.$$

So by the chain rule,

$$\frac{\partial u_b}{\partial x_i}\frac{\partial x_i}{\partial u_a} = \delta_{ab},$$

i.e. $M(f \circ g) = I$.

In the $n = 1$ case, it is the familiar result that $\mathrm{d}u/\mathrm{d}x = 1/(\mathrm{d}x/\mathrm{d}u)$.

Example. For $n = 2$, write $u_1 = \rho$, $u_2 = \varphi$ and let $x_1 = \rho\cos\varphi$ and $x_2 = \rho\sin\varphi$. Then the function used to convert between the coordinate systems is $g(u_1, u_2) = (u_1\cos u_2, u_1 \sin u_2)$

Then

$$M(g) = \begin{pmatrix} \partial x_1/\partial\rho & \partial x_1/\partial\varphi \\ \partial x_2/\partial\rho & \partial x_2/\partial\varphi \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\rho\sin\varphi \\ \sin\varphi & \rho\cos\varphi \end{pmatrix}$$

We can invert the relations between (x_1, x_2) and (ρ, φ) to obtain

$$\varphi = \tan^{-1}\frac{x_2}{x_1}$$
$$\rho = \sqrt{x_1^2 + x_2^2}$$

We can calculate

$$M(f) = \begin{pmatrix} \partial\rho/\partial x_1 & \partial\rho/\partial x_2 \\ \partial\varphi/\partial x_1 & \partial\varphi/\partial x_2 \end{pmatrix} = M(g)^{-1}.$$

These matrices are known as Jacobian matrices, and their determinants are known as the Jacobians.

Note that

$$\det M(f) \det M(g) = 1.$$

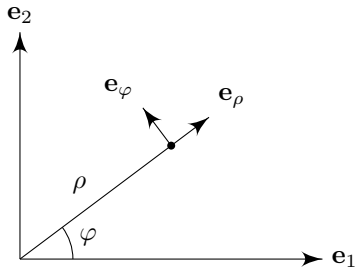
1.3 Coordinate systems

Now we can apply the results above the changes of coordinates on Euclidean space. Suppose x_i are the coordinates are Cartesian coordinates. Then we can define an arbitrary new coordinate system u_a in which each coordinate u_a is a function of $\mathbf{x}$. For example, we can define the plane polar coordinates ρ, φ by

$$x_1 = \rho\cos\varphi, \quad x_2 = \rho\sin\varphi.$$

However, note that ρ and φ are not components of a position vector, i.e. they are not the "coefficients" of basis vectors like $\mathbf{r} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$ are. But we can associate related basis vectors that point to directions of increasing ρ and φ, obtained by differentiating $\mathbf{r}$ with respect to the variables and then normalizing:

$$\mathbf{e}_\rho = \cos\varphi\,\mathbf{e}_1 + \sin\varphi\,\mathbf{e}_2, \quad \mathbf{e}_\varphi = -\sin\varphi\,\mathbf{e}_1 + \cos\varphi\,\mathbf{e}_2.$$



These are not "usual" basis vectors in the sense that these basis vectors vary with position and are undefined at the origin. However, they are still very useful when dealing with systems with rotational symmetry.

In three dimensions, we have cylindrical polars and spherical polars.

Cylindrical polars	Spherical polars
Conversion formulae	
$x_1 = \rho\cos\varphi$	$x_1 = r\sin\theta\cos\varphi$
$x_2 = \rho\sin\varphi$	$x_2 = r\sin\theta\sin\varphi$
$x_3 = z$	$x_3 = r\cos\theta$
Basis vectors	
$\mathbf{e}_\rho = (\cos\varphi, \sin\varphi, 0)$	$\mathbf{e}_r = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$
$\mathbf{e}_\varphi = (-\sin\varphi, \cos\varphi, 0)$	$\mathbf{e}_\varphi = (-\sin\varphi, \cos\varphi, 0)$
$\mathbf{e}_z = (0, 0, 1)$	$\mathbf{e}_\theta = (\cos\theta\cos\varphi, \cos\theta\sin\varphi, -\sin\theta)$

2 Curves and Line

2.1 Parametrised curves, lengths and arc length

There are many ways we can described a curve. We can, say, describe it by a equation that the points on the curve satisfy. For example, a circle can be described by $x^2 + y^2 = 1$. However, this is not a good way to do so, as it is rather difficult to work with. It is also often difficult to find a closed form like this for a curve.

Instead, we can imagine the curve to be specified by a particle moving along the path. So it is represented by a function $\mathbf{x} : \mathbb{R} \to \mathbb{R}^n$, and the curve itself is the image of the function. This is known as a *parametrisation* of a curve. In addition to simplified notation, this also has the benefit of giving the curve an *orientation*.

Definition (Parametrisation of curve)**.** Given a curve C in $\mathbb{R}^n$, a *parametrisation* of it is a continuous and invertible function $\mathbf{r} : D \to \mathbb{R}^n$ for some $D \subseteq \mathbb{R}$ whose image is C.

$\mathbf{r}'(u)$ is a vector tangent to the curve at each point. A parametrization is *regular* if $\mathbf{r}'(u) \neq 0$ for all u.

Clearly, a curve can have many different parametrizations.

Example. The curve

$$\frac{1}{4}x^2 + y^2 = 1, \quad y \geq 0, \quad z = 3.$$

can be parametrised by $2\cos u\hat{\mathbf{i}} + \sin u\hat{\mathbf{j}} + 3\hat{\mathbf{k}}$

If we change u (and hence $\mathbf{r}$) by a small amount, then the distance $|\delta\mathbf{r}|$ is roughly equal to the change in arclength δs. So $\delta s = |\delta\mathbf{r}| + o(\delta\mathbf{r})$. Then we have

Proposition. Let s denote the arclength of a curve $\mathbf{r}(u)$. Then

$$\frac{\mathrm{d}s}{\mathrm{d}u} = \pm\left|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}u}\right| = \pm|\mathbf{r}'(u)|$$

with the sign depending on whether it is in the direction of increasing or decreasing arclength.

Example. Consider a helix described by $\mathbf{r}(u) = (3\cos u, 3\sin u, 4u)$. Then

$$\mathbf{r}'(u) = (-3\sin u, 3\cos u, 4)$$
$$\frac{\mathrm{d}s}{\mathrm{d}u} = |\mathbf{r}'(u)| = \sqrt{3^2 + 4^2} = 5$$

So $s = 5u$. i.e. the arclength from $\mathbf{r}(0)$ and $\mathbf{r}(u)$ is $s = 5u$.

We can change parametrisation of $\mathbf{r}$ by taking an invertible smooth function $u \mapsto \tilde{u}$, and have a new parametrization $\mathbf{r}(\tilde{u}) = \mathbf{r}(\tilde{u}(u))$. Then by the chain rule,

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}u} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\tilde{u}} \times \frac{\mathrm{d}\tilde{u}}{\mathrm{d}u}$$
$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\tilde{u}} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}u} \Big/ \frac{\mathrm{d}\tilde{u}}{\mathrm{d}u}$$

It is often convenient to use the arclength s as the parameter. Then the tangent vector will always have unit length since the proposition above yields

$$|\mathbf{r}'(s)| = \frac{\mathrm{d}s}{\mathrm{d}s} = 1.$$

We call $\mathrm{d}s$ the scalar line element, which will be used when we consider integrals.

Definition (Scalar line element)**.** The *scalar line element* of C is $\mathrm{d}s$.

Proposition. $\mathrm{d}s = \pm|\mathbf{r}'(u)|\mathrm{d}u$

2.2 Line integrals of vector fields

Definition (Line integral)**.** The *line integral* of a smooth vector field $\mathbf{F}(\mathbf{r})$ along a path C parametrised by $\mathbf{r}(u)$ along the direction (orientation) $\mathbf{r}(\alpha) \to \mathbf{r}(\beta)$ is

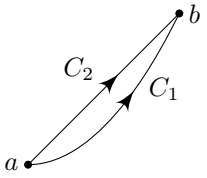
$$\int_C \mathbf{F}(\mathbf{r}) \cdot \mathrm{d}\mathbf{r} = \int_\alpha^\beta \mathbf{F}(\mathbf{r}(u)) \cdot \mathbf{r}'(u)\ \mathrm{d}u.$$

We say $\mathrm{d}\mathbf{r} = \mathbf{r}'(u)\mathrm{d}u$ is the *line element* on C. Note that the upper and lower limits of the integral are the end point and start point respectively, and β is not necessarily larger than α.

For example, we may be moving a particle from $\mathbf{a}$ to $\mathbf{b}$ along a curve C under a force field $\mathbf{F}$. Then we may divide the curve into many small segments $\delta\mathbf{r}$. Then for each segment, the force experienced is $\mathbf{F}(\mathbf{r})$ and the work done is $\mathbf{F}(\mathbf{r}) \cdot \delta\mathbf{r}$. Then the total work done across the curve is

$$W = \int_C \mathbf{F}(\mathbf{r}) \cdot \mathrm{d}\mathbf{r}.$$

Example. Take $\mathbf{F}(\mathbf{r}) = (xe^y, z^2, xy)$ and we want to find the line integral from $\mathbf{a} = (0, 0, 0)$ to $\mathbf{b} = (1, 1, 1)$.



We first integrate along the curve $C_1 : \mathbf{r}(u) = (u, u^2, u^3)$. Then $\mathbf{r}'(u) = (1, 2u, 3u^2)$, and $\mathbf{F}(\mathbf{r}(u)) = (ue^{u^2}, u^6, u^3)$. So

$$\begin{aligned}
\int_{C_1} \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= \int_0^1 \mathbf{F} \cdot \mathbf{r}'(u)\ \mathrm{d}u \\
&= \int_0^1 ue^{u^2} + 2u^7 + 3u^5\ \mathrm{d}u \\
&= \frac{e}{2} - \frac{1}{2} + \frac{1}{4} + \frac{1}{2} \\
&= \frac{e}{2} + \frac{1}{4}
\end{aligned}$$

Now we try to integrate along another curve $C_2 : \mathbf{r}(t) = (t, t, t)$. So $\mathbf{r}'(t) = (1, 1, 1)$.

$$\begin{aligned}\int_{C_2} \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= \int \mathbf{F} \cdot \mathbf{r}'(t) \mathrm{d}t \\ &= \int_0^1 te^t + 2t^2 \, \mathrm{d}t \\ &= \frac{5}{3}.\end{aligned}$$

We see that the line integral depends on the curve C in general, not just $\mathbf{a}, \mathbf{b}$.

We can also use the arclength s as the parameter. Since $\mathrm{d}\mathbf{r} = \mathbf{t} \, \mathrm{d}s$, with $\mathbf{t}$ being the unit tangent vector, we have

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_C \mathbf{F} \cdot \mathbf{t} \, \mathrm{d}s.$$

Note that we do not necessarily have to integrate $\mathbf{F} \cdot \mathbf{t}$ with respect to s. We can also integrate a scalar function as a function of s, $\int_C f(s) \, \mathrm{d}s$. By convention, this is calculated in the direction of increasing s. In particular, we have

$$\int_C 1 \, \mathrm{d}s = \text{length of C}.$$

Definition (Closed curve)**.** A *closed curve* is a curve with the same start and end point. The line integral along a closed curve is (sometimes) written as $\oint$ and is (sometimes) called the *circulation* of $\mathbf{F}$ around C.

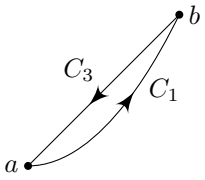
Sometimes we are not that lucky and our curve is not smooth. For example, the graph of an absolute value function is not smooth. However, often we can break it apart into many smaller segments, each of which is smooth. Alternatively, we can write the curve as a sum of smooth curves. We call these *piecewise smooth* curves.

Definition (Piecewise smooth curve)**.** A *piecewise smooth curve* is a curve $C = C_1 + C_2 + \cdots + C_n$ with all C_i smooth with regular parametrisations. The line integral over a piecewise smooth C is

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_{C_1} \mathbf{F} \cdot \mathrm{d}\mathbf{r} + \int_{C_2} \mathbf{F} \cdot \mathrm{d}\mathbf{r} + \cdots + \int_{C_n} \mathbf{F} \cdot \mathrm{d}\mathbf{r}.$$

Example. Take the example above, and let $C_3 = -C_2$. Then $C = C_1 + C_3$ is piecewise smooth but not smooth. Then

$$\begin{aligned}\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= \int_{C_1} \mathbf{F} \cdot \mathrm{d}\mathbf{r} + \int_{C_3} \mathbf{F} \cdot \mathrm{d}\mathbf{r} \\ &= \left(\frac{e}{2} + \frac{1}{4}\right) - \frac{5}{3} \\ &= -\frac{17}{12} + \frac{e}{2}.\end{aligned}$$



2.3 Gradients and Differentials

Recall that the line integral depends on the actual curve taken, and not just the end points. However, for some nice functions, the integral *does* depend on the end points only.

Theorem. If $\mathbf{F} = \nabla f(\mathbf{r})$, then

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = f(\mathbf{b}) - f(\mathbf{a}),$$

where $\mathbf{b}$ and $\mathbf{a}$ are the end points of the curve.

In particular, the line integral does *not* depend on the curve, but the end points only. This is the vector counterpart of the fundamental theorem of calculus. A special case is when C is a closed curve, then $\oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0$.

Proof. Let $\mathbf{r}(u)$ be any parametrization of the curve, and suppose $\mathbf{a} = \mathbf{r}(\alpha)$, $\mathbf{b} = \mathbf{r}(\beta)$. Then

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_C \nabla f \cdot \mathrm{d}\mathbf{r} = \int \nabla f \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}u}\, \mathrm{d}u.$$

So by the chain rule, this is equal to

$$\int_\alpha^\beta \frac{\mathrm{d}}{\mathrm{d}u}(f(\mathbf{r}(u)))\, \mathrm{d}u = [f(\mathbf{r}(u))]_\alpha^\beta = f(\mathbf{b}) - f(\mathbf{a}). \qquad \square$$

Definition (Conservative vector field)**.** If $\mathbf{F} = \nabla f$ for some f, the $\mathbf{F}$ is called a *conservative vector field*.

The name *conservative* comes from mechanics, where conservative vector fields represent conservative forces that conserve energy. This is since if the force is conservative, then the integral (i.e. work done) about a closed curve is 0, which means that we cannot gain energy after travelling around the loop.

It is convenient to treat differentials $\mathbf{F} \cdot \mathrm{d}\mathbf{r} = F_i \mathrm{d}x_i$ as if they were objects by themselves, which we can integrate along curves if we feel like doing so.

Then we can define

Definition (Exact differential)**.** A differential $\mathbf{F} \cdot \mathrm{d}\mathbf{r}$ is *exact* if there is an f such that $\mathbf{F} = \nabla f$. Then

$$\mathrm{d}f = \nabla f \cdot \mathrm{d}\mathbf{r} = \frac{\partial f}{\partial x_i} \mathrm{d}x_i.$$

To test if this holds, we can use the necessary condition

Proposition. If $\mathbf{F} = \nabla f$ for some f, then

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}.$$

This is because both are equal to $\partial^2 f / \partial x_i \partial x_j$.

For an exact differential, the result from the previous section reads

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_C \mathrm{d}f = f(\mathbf{b}) - f(\mathbf{a}).$$

Differentials can be manipulated using (for constant λ, μ):

Proposition.

$$\mathrm{d}(\lambda f + \mu g) = \lambda \mathrm{d}f + \mu \mathrm{d}g$$
$$\mathrm{d}(fg) = (\mathrm{d}f)g + f(\mathrm{d}g)$$

Using these, it may be possible to find f by inspection.

Example. Consider

$$\int_C 3x^2 y \sin z \, \mathrm{d}x + x^3 \sin z \, \mathrm{d}y + x^3 y \cos z \, \mathrm{d}z.$$

We see that if we integrate the first term with respect to x, we obtain $x^3 y \sin z$. We obtain the same thing if we integrate the second and third term. So this is equal to

$$\int_C \mathrm{d}(x^3 y \sin z) = [x^3 y \sin z]_{\mathbf{a}}^{\mathbf{b}}.$$

2.4 Work and potential energy

Definition (Work and potential energy)**.** If $\mathbf{F}(\mathbf{r})$ is a force, then $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ is the *work done* by the force along the curve C. It is the limit of a sum of terms $\mathbf{F}(\mathbf{r}) \cdot \delta\mathbf{r}$, i.e. the force along the direction of $\delta\mathbf{r}$.

Consider a point particle moving under $\mathbf{F}(\mathbf{r})$ according to Newton's second law: $\mathbf{F}(\mathbf{r}) = m\ddot{\mathbf{r}}$.

Since the kinetic energy is defined as

$$T(t) = \frac{1}{2} m \dot{\mathbf{r}}^2,$$

the rate of change of energy is

$$\frac{\mathrm{d}}{\mathrm{d}t} T(t) = m\dot{\mathbf{r}} \cdot \ddot{\mathbf{r}} = \mathbf{F} \cdot \dot{\mathbf{r}}.$$

Suppose the path of particle is a curve C from $\mathbf{a} = \mathbf{r}(\alpha)$ to $\mathbf{b} = \mathbf{r}(\beta)$, Then

$$T(\beta) - T(\alpha) = \int_\alpha^\beta \frac{\mathrm{d}T}{\mathrm{d}t} \, \mathrm{d}t = \int_\alpha^\beta \mathbf{F} \cdot \dot{\mathbf{r}} \, \mathrm{d}t = \int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}.$$

So the work done on the particle is the change in kinetic energy.

Definition (Potential energy)**.** Given a conservative force $\mathbf{F} = -\nabla V$, $V(\mathbf{x})$ is the *potential energy*. Then

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = V(\mathbf{a}) - V(\mathbf{b}).$$

Therefore, for a conservative force, we have $\mathbf{F} = \nabla V$, where $V(\mathbf{r})$ is the potential energy.

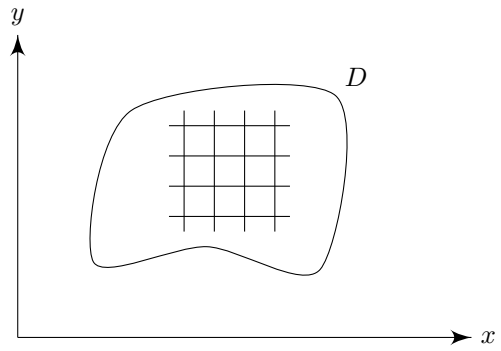
So the work done (gain in kinetic energy) is the loss in potential energy. So the total energy $T + V$ is conserved, i.e. constant during motion.

We see that energy is conserved for conservative forces. In fact, the converse is true — the energy is conserved only for conservative forces.

3 Integration in $\mathbb{R}^2$ and $\mathbb{R}^3$

3.1 Integrals over subsets of $\mathbb{R}^2$

Definition (Surface integral)**.** Let $D \subseteq \mathbb{R}^2$. Let $\mathbf{r} = (x, y)$ be in Cartesian coordinates. We can approximate D by N disjoint subsets of simple shapes, e.g. triangles, parallelograms. These shapes are labelled by I and have areas δA_i.



To integrate a function f over D, we would like to take the sum $\sum f(\mathbf{r}_i)\delta A_i$, and take the limit as $\delta A_i \to 0$. But we need a condition stronger than simply $\delta A_i \to 0$. We won't want the areas to grow into arbitrarily long yet thin strips whose area decreases to 0. So we say that we find an ℓ such that each area can be contained in a disc of diameter ℓ.

Then we take the limit as $\ell \to 0$, $N \to \infty$, and the union of the pieces tends to D. For a function $f(\mathbf{r})$, we define the *surface integral* as

$$\int_D f(\mathbf{r})\, \mathrm{d}A = \lim_{\ell \to 0} \sum_I f(\mathbf{r}_i)\delta A_i.$$

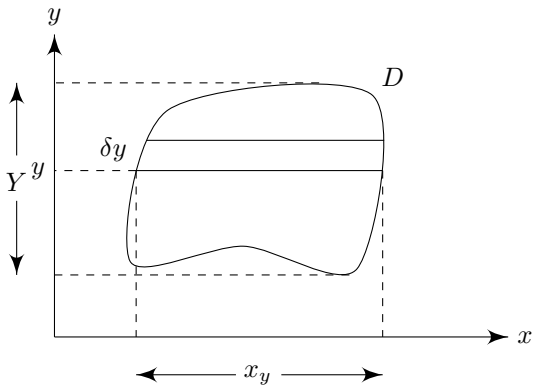
where $\mathbf{r}_i$ is some point within each subset A_i. The integral *exists* if the limit is well-defined (i.e. the same regardless of what A_i and $\mathbf{r}_i$ we choose before we take the limit) and exists.

If we take $f = 1$, then the surface integral is the area of D.

On the other hand, if we put $z = f(x, y)$ and plot out the surface $z = f(x, y)$, then the area integral is the volume under the surface.

The definition allows us to take the δA_i to be any weird shape we want. However, the sensible thing is clearly to take A_i to be rectangles.

We choose the small sets in the definition to be rectangles, each of size $\delta A_I = \delta x \delta y$. We sum over subsets in a narrow horizontal strip of height δy with y and δy held constant. Take the limit as $\delta x \to 0$. We get a contribution $\delta y \int_{x_y} f(y, x)\, \mathrm{d}x$ with range $x_y \in \{x : (x, y) \in D\}$.



We sum over all such strips and take $\delta y \to 0$, giving

Proposition.

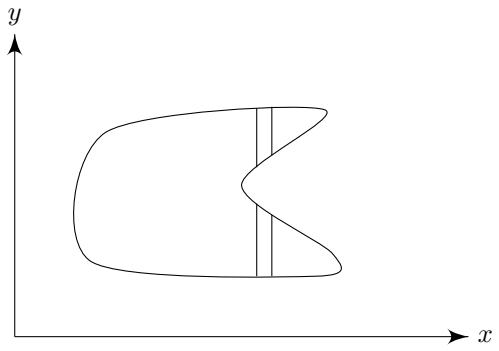
$$\int_D f(x,y)\,\mathrm{d}A = \int_Y \left(\int_{x_y} f(x,y)\,\mathrm{d}x \right) \mathrm{d}y.$$

with x_y ranging over $\{x : (x,y) \in D\}$.

Note that the range of the inner integral is given by a set x_y. This can be an interval, or many disconnected intervals, $x_y = [a_1, b_1] \cup [a_2, b_2]$. In this case,

$$\int_{x_y} f(x)\,\mathrm{d}x = \int_{a_1}^{b_1} f(x)\,\mathrm{d}x + \int_{a_2}^{b_2} f(x)\,\mathrm{d}x.$$

This is useful if we want to integrate over a concave area and we have disconnected vertical strips.



We could also do it the other way round, integrating over y first, and come up with the result

$$\int_D f(x,y)\,\mathrm{d}A = \int_X \left(\int_{y_x} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x.$$

Theorem (Fubini's theorem)**.** If f is a continuous function and D is a compact (i.e. closed and bounded) subset of $\mathbb{R}^2$, then

$$\iint f\,\mathrm{d}x\,\mathrm{d}y = \iint f\,\mathrm{d}y\,\mathrm{d}x.$$

While we have rather strict conditions for this theorem, it actually holds in many more cases, but those situations have to be checked manually.

Definition (Area element)**.** The *area element* is $\mathrm{d}A$.

Proposition. $\mathrm{d}A = \mathrm{d}x\ \mathrm{d}y$ in Cartesian coordinates.

Example. We integrate over the triangle bounded by $(0,0), (2,0)$ and $(0,1)$. We want to integrate the function $f(x,y) = x^2y$ over the area. So

$$\begin{aligned}\int_D f(xy)\ \mathrm{d}A &= \int_0^1 \left(\int_0^{2-2y} x^2y\ \mathrm{d}x\right)\ \mathrm{d}y\\ &= \int_0^1 y\left[\frac{x^3}{3}\right]_0^{2-2y}\ \mathrm{d}y\\ &= \frac{8}{3}\int_0^1 y(1-y)^3\ \mathrm{d}y\\ &= \frac{2}{15}\end{aligned}$$

We can integrate it the other way round:

$$\begin{aligned}\int_D x^2y\ \mathrm{d}A &= \int_0^2 \int_0^{1-x/2} x^2y\ \mathrm{d}y\ \mathrm{d}x\\ &= \int_0^2 x^2\left[\frac{1}{2}y^2\right]_0^{1-x/2}\ \mathrm{d}x\\ &= \int_0^2 \frac{x^2}{2}\left(1-\frac{x}{2}\right)^2\ \mathrm{d}x\\ &= \frac{2}{15}\end{aligned}$$

Since it doesn't matter whether we integrate x first or y first, if we find it difficult to integrate one way, we can try doing it the other way and see if it is easier.

While this integral is tedious in general, there is a special case where it is substantially easier.

Definition (Separable function)**.** A function $f(x,y)$ is separable if it can be written as $f(x,y) = h(y)g(x)$.

Proposition. Take separable $f(x,y) = h(y)g(x)$ and D be a rectangle $\{(x,y) : a \le x \le b, c \le y \le d\}$. Then

$$\int_D f(x,y)\ \mathrm{d}x\ \mathrm{d}y = \left(\int_a^b g(x)\ \mathrm{d}x\right)\left(\int_c^d h(y)\ \mathrm{d}y\right)$$

3.2 Change of variables for an integral in $\mathbb{R}^2$

Proposition. Suppose we have a change of variables $(x,y) \leftrightarrow (u,v)$ that is smooth and invertible, with regions D, D' in one-to-one correspondence. Then

$$\int_D f(x,y)\ \mathrm{d}x\ \mathrm{d}y = \int_{D'} f(x(u,v), y(u,v))|J|\ \mathrm{d}u\ \mathrm{d}v,$$

where

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

is the Jacobian. In other words,

$$\mathrm{d}x\, \mathrm{d}y = |J|\, \mathrm{d}u\, \mathrm{d}v.$$

Proof. Since we are writing $(x(u, v), y(u, v))$, we are actually transforming from (u, v) to (x, y) and not the other way round.

Suppose we start with an area $\delta A' = \delta u \delta v$ in the (u, v) plane. Then by Taylors' theorem, we have

$$\delta x = x(u + \delta u, v + \delta v) - x(u, v) \approx \frac{\partial x}{\partial u}\delta u + \frac{\partial x}{\partial v}\delta v.$$

We have a similar expression for δy and we obtain

$$\begin{pmatrix} \delta x \\ \delta y \end{pmatrix} \approx \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} \begin{pmatrix} \delta u \\ \delta v \end{pmatrix}$$

Recall from Vectors and Matrices that the determinant of the matrix is how much it scales up an area. So the area formed by δx and δy is $|J|$ times the area formed by δu and δv. Hence

$$\mathrm{d}x\, \mathrm{d}y = |J|\, \mathrm{d}u\, \mathrm{d}v. \qquad \square$$

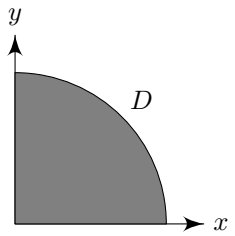
Example. We transform from (x, y) to (ρ, φ) with

$$x = \rho \cos \varphi$$
$$y = \rho \sin \varphi$$

We have previously calculated that $|J| = \rho$. So

$$\mathrm{d}A = \rho\, \mathrm{d}\rho\, \mathrm{d}\varphi.$$

Suppose we want to integrate a function over a quarter area D of radius R.



Let the function to be integrated be $f = \exp(-(x^2 + y^2)/2) = \exp(-\rho^2/2)$. Then

$$\begin{aligned} \int f\, \mathrm{d}A &= \int f \rho\, \mathrm{d}\rho\, \mathrm{d}\varphi \\ &= \int_{\rho=0}^{R} \left(\int_{\varphi=0}^{\pi/2} e^{-\rho^2/2} \rho\, \mathrm{d}\varphi \right) \delta\rho \end{aligned}$$

Note that in polar coordinates, we are integrating over a rectangle and the function is separable. So this is equal to

$$\begin{aligned} &= \left[-e^{-\rho^2/2}\right]_0^R [\varphi]_0^{\pi/2} \\ &= \frac{\pi}{2}\left(1 - e^{-R^2/2}\right). \end{aligned} \tag{$*$}$$

Note that the integral exists as $R \to \infty$.

Now we take the case of $x, y \to \infty$ and consider the original integral.

$$\begin{aligned} \int_D f \, \mathrm{d}A &= \int_{x=0}^{\infty}\int_{y=0}^{\infty} e^{-(x^2+y^2)/2} \, \mathrm{d}x \, \mathrm{d}y \\ &= \left(\int_0^{\infty} e^{-x^2/2} \, \mathrm{d}x\right)\left(\int_0^{\infty} e^{-y^2/2} \, \mathrm{d}y\right) \\ &= \frac{\pi}{2} \end{aligned}$$

where the last line is from (*). So each of the two integrals must be $\sqrt{\pi/2}$, i.e.

$$\int_0^{\infty} e^{-x^2/2} \, \mathrm{d}x = \sqrt{\frac{\pi}{2}}.$$

3.3 Generalization to $\mathbb{R}^3$

We will do exactly the same thing as we just did, but with one more dimension:

Definition (Volume integral)**.** Consider a volume $V \subseteq \mathbb{R}^3$ with position vector $\mathbf{r} = (x, y, z)$. We approximate V by N small disjoint subsets of some simple shape (e.g. cuboids) labelled by I, volume δV_I, contained within a solid sphere of diameter ℓ.

Assume that as $\ell \to 0$ and $N \to \infty$, the union of the small subsets tend to V. Then

$$\int_V f(\mathbf{r}) \, \mathrm{d}V = \lim_{\ell \to 0} \sum_I f(\mathbf{r}_I^*)\delta V_I,$$

where $\mathbf{r}_I^*$ is any chosen point in each small subset.

To evaluate this, we can take $\delta V_I = \delta x \delta y \delta z$, and take $\delta x \to 0$, $\delta y \to 0$ and δz in some order. For example,

$$\int_V f(\mathbf{r}) \, \mathrm{d}v = \int_D \left(\int_{Z_{xy}} f(x, y, z) \, \mathrm{d}z\right) \, \mathrm{d}x \, \mathrm{d}y.$$

So we integrate $f(x, y, z)$ over z at each point (x, y), then take the integral of that over the area containing all required (x, y).

Alternatively, we can take the area integral first, and have

$$\int_V f(\mathbf{r}) \, \mathrm{d}V = \int_z \left(\int_{D_Z} f(x, y, z) \, \mathrm{d}x \, \mathrm{d}y\right) \, \mathrm{d}z.$$

Again, if we take $f = 1$, then we obtain the volume of V.

Often, $f(\mathbf{r})$ is the density of some quantity, and is usually denoted by ρ. For example, we might have mass density, charge density, or probability density. $\rho(\mathbf{r})\delta V$ is then the amount of quantity in a small volume δV at $\mathbf{r}$. Then $\int_V \rho(\mathbf{r}) \, \mathrm{d}V$ is the total amount of quantity in V.

Definition (Volume element)**.** The *volume element* is $\mathrm{d}V$.

Proposition. $\mathrm{d}V = \mathrm{d}x\ \mathrm{d}y\ \mathrm{d}z$.

We can change variables by some smooth, invertible transformation $(x, y, z) \mapsto (u, v, w)$. Then

Proposition.

$$\int_V f\ \mathrm{d}x\ \mathrm{d}y\ \mathrm{d}z = \int_V f|J|\ \mathrm{d}u\ \mathrm{d}v\ \mathrm{d}w,$$

with

$$J = \frac{\partial(x, y, z)}{\partial(u, v, w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix}$$

Proposition. In cylindrical coordinates,

$$\mathrm{d}V = \rho\ \mathrm{d}\rho\ \mathrm{d}\varphi\ \mathrm{d}z.$$

In spherical coordinates

$$\mathrm{d}V = r^2 \sin\theta\ \mathrm{d}r\ \mathrm{d}\theta\ \mathrm{d}\varphi.$$

Proof. Loads of algebra. □

Example. Suppose $f(\mathbf{r})$ is spherically symmetric and V is a sphere of radius a centered on the origin. Then

$$\begin{aligned} \int_V f\ \mathrm{d}V &= \int_{r=0}^{a} \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} f(r) r^2 \sin\theta\ \mathrm{d}r\ \mathrm{d}\theta\ \mathrm{d}\varphi \\ &= \int_0^a \mathrm{d}r \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\varphi\ r^2 f(r) \sin\theta \\ &= \int_0^a r^2 f(r) \mathrm{d}r \Big[-\cos\theta \Big]_0^\pi \Big[\varphi \Big]_0^{2\pi} \\ &= 4\pi \int_0^a f(r) r^2\ \mathrm{d}r. \end{aligned}$$

where we separated the integral into three parts as in the area integrals.

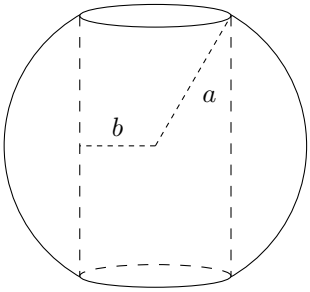
Note that in the second line, we rewrote the integrals to write the differentials next to the integral sign. This is simply a different notation that saves us from writing $r = 0$ etc. in the limits of the integrals.

This is a useful general result. We understand it as the sum of spherical shells of thickness δr and volume $4\pi r^2 \delta r$.

If we take $f = 1$, then we have the familiar result that the volume of a sphere is $\frac{4}{3}\pi a^3$.

Example. Consider a volume within a sphere of radius a with a cylinder of radius b ($b < a$) removed. The region is defined as

$$\begin{aligned} x^2 + y^2 + z^2 &\leq a^2 \\ x^2 + y^2 &\geq b^2. \end{aligned}$$



We use cylindrical coordinates. The second criteria gives

$$b \le \rho \le a.$$

For the $x^2 + y^2 + z^2 \le a^2$ criterion, we have

$$-\sqrt{a^2 - \rho^2} \le z \le \sqrt{a^2 - \rho^2}.$$

So the volume is

$$\begin{aligned}\int_V \mathrm{d}V &= \int_b^a \mathrm{d}\rho \int_0^{2\pi} \mathrm{d}\varphi \int_{-\sqrt{a^2-\rho^2}}^{\sqrt{a^2-\rho^2}} \mathrm{d}z\, \rho \\ &= 2\pi \int_b^a 2\rho\sqrt{a^2 - \rho^2}\, \mathrm{d}\rho \\ &= 2\pi \left[\frac{2}{3}(a^2 - \rho^2)^{3/2}\right]_b^a \\ &= \frac{4}{3}\pi(a^2 - b^2)^{3/2}.\end{aligned}$$

Example. Suppose the density of electric charge is $\rho(\mathbf{r}) = \rho_0 \frac{z}{a}$ in a hemisphere H of radius a, with $z \ge 0$. What is the total charge of H?

We use spherical polars. So

$$r \le a, \quad 0 \le \varphi \le 2\pi, \quad 0 \le \theta \le \frac{\pi}{2}.$$

We have

$$\rho(\mathbf{r}) = \frac{\rho_0}{a} r\cos\theta.$$

The total charge Q in H is

$$\begin{aligned}\int_H \rho\, \mathrm{d}V &= \int_0^a \mathrm{d}r \int_0^{\pi/2} \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\varphi\, \frac{\rho_0}{a} r\cos\theta r^2 \sin\theta \\ &= \frac{\rho_0}{a}\int_0^a r^3\, \mathrm{d}r \int_0^{\pi/2} \sin\theta\cos\theta\, \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\varphi \\ &= \frac{\rho_0}{a}\left[\frac{r^4}{4}\right]_0^a \left[\frac{1}{2}\sin^2\theta\right]_0^{\pi/2} [\varphi]_0^{2\pi} \\ &= \frac{\rho_0 \pi a^3}{4}.\end{aligned}$$

3.4 Further generalizations

Integration in $\mathbb{R}^n$

Similar to the above, $\int_D f(x_1, x_2, \cdots x_n)\ \mathrm{d}x_1\ \mathrm{d}x_2 \cdots\ \mathrm{d}x_n$ is simply the integration over an n-dimensional volume. The change of variable formula is

Proposition.

$$\int_D f(x_1, x_2, \cdots x_n)\ \mathrm{d}x_1\ \mathrm{d}x_2 \cdots\ \mathrm{d}x_n = \int_{D'} f(\{x_i(\mathbf{u})\})|J|\ \mathrm{d}u_1\ \mathrm{d}u_2 \cdots\ \mathrm{d}u_n.$$

Change of variables for $n = 1$

In the $n = 1$ case, the Jacobian is $\frac{\mathrm{d}x}{\mathrm{d}u}$. However, we use the following formula for change of variables:

$$\int_D f(x)\ \mathrm{d}x = \int_{D'} f(x(u)) \left|\frac{\mathrm{d}x}{\mathrm{d}u}\right|\ \mathrm{d}u.$$

We introduce the modulus because of our natural convention about integrating over D and D'. If $D = [a, b]$ with $a < b$, we write $\int_a^b$. But if $a \mapsto \alpha$ and $b \mapsto \beta$, but $\alpha > \beta$, we would like to write $\int_\beta^\alpha$ instead, so we introduce the modulus in the 1D case.

To show that the modulus is the right thing to do, we check case by case: If $a < b$ and $\alpha < \beta$, then $\frac{\mathrm{d}x}{\mathrm{d}u}$ is positive, and we have, as expected

$$\int_a^b f(x)\ \mathrm{d}x = \int_\alpha^\beta f(u)\frac{\mathrm{d}x}{\mathrm{d}u}\ \mathrm{d}u.$$

If $\alpha > \beta$, then $\frac{\mathrm{d}x}{\mathrm{d}u}$ is negative. So

$$\int_a^b f(x)\ \mathrm{d}x = \int_\alpha^\beta f(u)\frac{\mathrm{d}x}{\mathrm{d}u}\ \mathrm{d}u = -\int_\beta^\alpha f(u)\frac{\mathrm{d}x}{\mathrm{d}u}\ \mathrm{d}u.$$

By taking the absolute value of $\frac{\mathrm{d}x}{\mathrm{d}u}$, we ensure that we always have the numerically smaller bound as the lower bound.

This is not easily generalized to higher dimensions, so we don't employ the same trick in other cases.

Vector-valued integrals

We can define $\int_V \mathbf{F}(\mathbf{r})\ \mathrm{d}V$ in a similar way to $\int_V f(\mathbf{r})\ \mathrm{d}V$as the limit of a sum over small contributions of volume. In practice, we integrate them componentwise. If

$$\mathbf{F}(\mathbf{r}) = F_i(\mathbf{r})\mathbf{e}_i,$$

then

$$\int_V \mathbf{F}(\mathbf{r})\ \mathrm{d}V = \int_V (F_i(\mathbf{r})\ \mathrm{d}V)\mathbf{e}_i.$$

For example, if a mass has density $\rho(\mathbf{r})$, then its mass is

$$M = \int_V \rho(\mathbf{r})\ \mathrm{d}V$$

and its center of mass is

$$\mathbf{R} = \frac{1}{M}\int_V \mathbf{r}\rho(\mathbf{r})\,\mathrm{d}V.$$

Example. Consider a solid hemisphere H with $r \le a$, $z \ge 0$ with uniform density ρ. The mass is

$$M = \int_H \rho\,\mathrm{d}V = \frac{2}{3}\pi a^3\rho.$$

Now suppose that $\mathbf{R} = (X, Y, Z)$. By symmetry, we expect $X = Y = 0$. We can find this formally by

$$\begin{aligned} X &= \frac{1}{M}\int_H x\rho\,\mathrm{d}V \\ &= \frac{\rho}{M}\int_0^a\int_0^{\pi/2}\int_0^{2\pi} xr^2\sin\theta\,\mathrm{d}\varphi\,\mathrm{d}\theta\,\mathrm{d}r \\ &= \frac{\rho}{M}\int_0^a r^3\,\mathrm{d}r \times \int_0^{\pi/2}\sin^2\theta\,\mathrm{d}\theta \times \int_0^{2\pi}\cos\varphi\,\mathrm{d}\varphi \\ &= 0 \end{aligned}$$

as expected. Note that it evaluates to 0 because the integral of cos from 0 to 2π is 0. Similarly, we obtain $Y = 0$.

Finally, we find Z.

$$\begin{aligned} Z &= \frac{\rho}{M}\int_0^a r^3\,\mathrm{d}r\int_0^{\pi/2}\sin\theta\cos\theta\,\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\varphi \\ &= \frac{r}{M}\left[\frac{a^4}{4}\right]\left[\frac{1}{2}\sin^2\theta\right]_0^{\pi/2} 2\pi \\ &= \frac{3a}{8}. \end{aligned}$$

So $\mathbf{R} = (0, 0, 3a/8)$.

4 Surfaces and surface integrals

4.1 Surfaces and Normal

So far, we have learnt how to do calculus with regions of the plane or space. What we would like to do now is to study surfaces in $\mathbb{R}^3$. The first thing to figure out is how to specify surfaces. One way to specify a surface is to use an equation. We let f be a smooth function on $\mathbb{R}^3$, and c be a constant. Then $f(\mathbf{r}) = c$ defines a smooth surface (e.g. $x^2 + y^2 + z^2 = 1$ denotes the unit sphere).

Now consider any curve $\mathbf{r}(u)$ on S. Then by the chain rule, if we differentiate $f(\mathbf{r}) = c$ with respect to u, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}u}[f(\mathbf{r}(u))] = \nabla f \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}u} = 0.$$

This means that ∇f is always perpendicular to $\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}u}$. Since $\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}u}$ is the tangent to the curve, ∇f is perpendicular to the tangent. Since this is true for *any* curve $\mathbf{r}(u)$, ∇f is perpendicular to any tangent of the surface. Therefore

Proposition. ∇f is the normal to the surface $f(\mathbf{r}) = c$.

Example.

(i) Take the sphere $f(\mathbf{r}) = x^2 + y^2 + z^2 = c$ for $c > 0$. Then $\nabla f = 2(x, y, z) = 2\mathbf{r}$, which is clearly normal to the sphere.

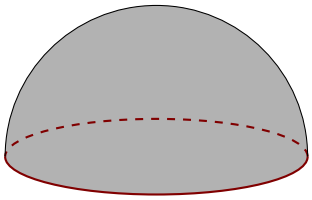
(ii) Take $f(\mathbf{r}) = x^2 + y^2 - z^2 = c$, which is a hyperboloid. Then $\nabla f = 2(x, y, -z)$.

In the special case where $c = 0$, we have a double cone, with a singular apex $\mathbf{0}$. Here $\nabla f = \mathbf{0}$, and we cannot find a meaningful direction of normal.

Definition (Boundary)**.** A surface S can be defined to have a *boundary* ∂S consisting of a piecewise smooth curve. If we define S as in the above examples but with the additional restriction $z \geq 0$, then ∂S is the circle $x^2 + y^2 = c$, $z = 0$.

A surface is *bounded* if it can be contained in a solid sphere, *unbounded* otherwise. A bounded surface with no boundary is called *closed* (e.g. sphere).

Example.



The boundary of a hemisphere is a circle (drawn in red).

Definition (Orientable surface)**.** At each point, there is a unit normal $\mathbf{n}$ that's unique up to a sign.

If we can find a consistent choice of $\mathbf{n}$ that varies smoothly across S, then we say S is *orientable*, and the choice of sign of $\mathbf{n}$ is called the *orientation* of the surface.

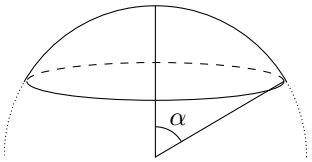
Most surfaces we encounter are orientable. For example, for a sphere, we can declare that the normal should always point outwards. A notable example of a non-orientable surface is the Möbius strip (or Klein bottle).

For simple cases, we can describe the orientation as "inward" and "outward".

4.2 Parametrized surfaces and area

However, specifying a surface by an equation $f(\mathbf{r}) = c$ is often not too helpful. What we would like is to put some coordinate system onto the surface, so that we can label each point by a pair of numbers (u, v), just like how we label points in the x, y-plane by (x, y). We write $\mathbf{r}(u, v)$ for the point labelled by (u, v).

Example. Let S be part of a sphere of radius a with $0 \leq \theta \leq \alpha$.



We can then label the points on the spheres by the angles θ, φ, with

$$\mathbf{r}(\theta, \varphi) = (a\cos\varphi\sin\theta, a\sin\theta\sin\varphi, a\cos\theta) = a\mathbf{e}_r.$$

We restrict the values of θ, φ by $0 \leq \theta \leq \alpha$, $0 \leq \varphi \leq 2\pi$, so that each point is only covered once.

Note that to specify a surface, in addition to the function $\mathbf{r}$, we also have to specify what values of (u, v) we are allowed to take. This corresponds to a region D of allowed values of u and v. When we do integrals with these surfaces, these will become the bounds of integration.

When we have such a parametrization $\mathbf{r}$, we would want to make sure this indeed gives us a two-dimensional surface. For example, the following two parametrizations would both be bad:

$$\mathbf{r}(u, v) = u, \quad \mathbf{r}(u, v) = u + v.$$

The idea is that $\mathbf{r}$ has to depend on both u and v, and in "different ways". More precisely, when we vary the coordinates (u, v), the point $\mathbf{r}$ will change accordingly. By the chain rule, this is given by

$$\delta\mathbf{r} = \frac{\partial\mathbf{r}}{\partial u}\delta u + \frac{\partial\mathbf{r}}{\partial v}\delta v + o(\delta u, \delta v).$$

Then $\frac{\partial\mathbf{r}}{\delta u}$ and $\frac{\partial\mathbf{r}}{\partial v}$ are tangent vectors to curves on S with v and u constant respectively. What we want is for them to point in different directions.

Definition (Regular parametrization)**.** A parametrization is *regular* if for all u, v,

$$\frac{\partial\mathbf{r}}{\partial u} \times \frac{\partial\mathbf{r}}{\partial v} \neq \mathbf{0},$$

i.e. there are always two independent tangent directions.

The parametrizations we use will all be regular.

Given a surface, how could we, say, find its area? We can use our parametrization. Suppose points on the surface are given by $\mathbf{r}(u, v)$ for $(u, v) \in D$. If we want to find the area of D itself, we would simply integrate

$$\int_D \mathrm{d}u\, \mathrm{d}v.$$

However, we are just using u and v as arbitrary labels for points in the surface, and one unit of area in D does not correspond to one unit of area in S. Instead, suppose we produce a small rectangle in D by changing u and v by small $\delta u, \delta v$. In D, this corresponds to a rectangle with vertices $(u, v), (u + \delta u, v), (u, v + \delta v), (u + \delta u, v + \delta v)$, and spans an area $\delta u \delta v$. In the surface S, these small changes $\delta u, \delta v$ correspond to changes $\frac{\partial \mathbf{r}}{\partial u}\delta u$ and $\frac{\partial \mathbf{r}}{\partial v}\delta v$, and these span a *vector* area of

$$\delta \mathbf{S} = \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \delta u \delta v = \mathbf{n}\, \delta S.$$

Note that the order of u, v gives the choice of the sign of the unit normal.

The actual area is then given by

$$\delta S = \left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right| \delta u\, \delta v.$$

Making these into differentials instead of deltas, we have

Proposition. The *vector area element* is

$$\mathrm{d}\mathbf{S} = \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\, \mathrm{d}u\, \mathrm{d}v.$$

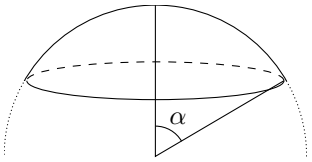
The *scalar area element* is

$$\mathrm{d}S = \left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right| \mathrm{d}u\, \mathrm{d}v.$$

By summing and taking limits, the area of S is

$$\int_S \mathrm{d}S = \int_D \left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right| \mathrm{d}u\, \mathrm{d}v.$$

Example. Consider again the part of the sphere of radius a with $0 \leq \theta \leq \alpha$.



Then we have

$$\mathbf{r}(\theta, \varphi) = (a \cos\varphi \sin\theta, a \sin\theta \sin\varphi, a\cos\theta) = a\mathbf{e}_r.$$

So we find

$$\frac{\partial \mathbf{r}}{\partial \theta} = a\mathbf{e}_\theta.$$

Similarly, we have

$$\frac{\partial \mathbf{r}}{\partial \varphi} = a \sin\theta \mathbf{e}_\varphi.$$

Then

$$\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi} = a^2 \sin\theta\, \mathbf{e}_r.$$

So

$$\mathrm{d}S = a^2 \sin\theta \,\mathrm{d}\theta \,\mathrm{d}\varphi.$$

Our bounds are $0 \leq \theta \leq \alpha$, $0 \leq \varphi \leq 2\pi$.

Then the area is

$$\int_0^{2\pi} \int_0^{\alpha} a^2 \sin\theta \,\mathrm{d}\theta \,\mathrm{d}\varphi = 2\pi a^2(1 - \cos\alpha).$$

4.3 Surface integral of vector fields

Just computing the area of a surface would be boring. Suppose we have a surface S parametrized by $\mathbf{r}(u, v)$, where (u, v) takes values in D. We would like to ask how much "stuff" is passing through S, where the flow of stuff is given by a vector field $\mathbf{F}(\mathbf{r})$.

We might attempt to use the integral

$$\int_D |\mathbf{F}| \,\mathrm{d}S.$$

However, this doesn't work. For example, if all the flow is tangential to the surface, then nothing is really passing through the surface, but $|\mathbf{F}|$ is non-zero, so we get a non-zero integral. Instead, what we should do is to consider the component of $\mathbf{F}$ that is *normal* to the surface S, i.e. parallel to its normal.

Definition (Surface integral)**.** The *surface integral* or *flux* of a vector field $\mathbf{F}(\mathbf{r})$ over S is defined by

$$\int_S \mathbf{F}(\mathbf{r}) \cdot \mathrm{d}\mathbf{S} = \int_S \mathbf{F}(\mathbf{r}) \cdot \mathbf{n} \,\mathrm{d}S = \int_D \mathbf{F}(\mathbf{r}(u, v)) \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right) \mathrm{d}u \,\mathrm{d}v.$$

Intuitively, this is the total amount of $\mathbf{F}$ passing through S. For example, if $\mathbf{F}$ is the electric field, the flux is the amount of electric field passing through a surface.

For a given orientation, the integral $\int \mathbf{F} \cdot \mathrm{d}\mathbf{S}$ is independent of the parametrization. Changing orientation is equivalent to changing the sign of $\mathbf{n}$, which is in turn equivalent to changing the order of u and v in the definition of S, which is also equivalent to changing the sign of the flux integral.

Example. Consider a sphere of radius a, $\mathbf{r}(\theta, \varphi)$. Then

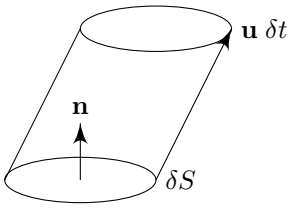
$$\frac{\partial \mathbf{r}}{\partial \theta} = a\mathbf{e}_\theta, \quad \frac{\partial \mathbf{r}}{\partial \varphi} = a \sin\theta \mathbf{e}_\varphi.$$

The vector area element is

$$\mathrm{d}\mathbf{S} = a^2 \sin\theta \mathbf{e}_r \,\mathrm{d}\theta \,\mathrm{d}\varphi,$$

taking the outward normal $\mathbf{n} = \mathbf{e}_r = \mathbf{r}/a$.

Suppose we want to calculate the fluid flux through the surface. The *velocity field* $\mathbf{u}(\mathbf{r})$ of a fluid gives the motion of a small volume of fluid $\mathbf{r}$. Assume that $\mathbf{u}$ depends smoothly on $\mathbf{r}$ (and t). For any small area δS, on a surface S, the volume of fluid crossing it in time δt is $\mathbf{u} \cdot \delta\mathbf{S}\ \delta t$.



So the amount of flow of $\mathbf{u}$ over at time δt through S is

$$\delta t \int_S \mathbf{u} \cdot \mathrm{d}\mathbf{S}.$$

So $\int_S \mathbf{u} \cdot \mathrm{d}\mathbf{S}$ is the *rate* of volume crossing S.

For example, let $\mathbf{u} = (-x, 0, z)$ and S be the section of a sphere of radius a with $0 \le \varphi \le 2\pi$ and $0 \le \theta \le \alpha$. Then

$$\mathrm{d}\mathbf{S} = a^2 \sin\theta \mathbf{n}\ \mathrm{d}\varphi\ \mathrm{d}\theta,$$

with

$$\mathbf{n} = \frac{\mathbf{r}}{a} = \frac{1}{a}(x, y, z).$$

So

$$\mathbf{n} \cdot \mathbf{u} = \frac{1}{a}(-x^2 + z^2) = a(-\sin^2\theta\cos^2\varphi + \cos^2\theta).$$

Therefore

$$\begin{aligned}
\int_S \mathbf{u} \cdot \mathrm{d}\mathbf{S} &= \int_0^\alpha \int_0^{2\pi} a^3 \sin\theta[(\cos^2\theta - 1)\cos^2\varphi + \cos^2\theta]\ \mathrm{d}\varphi\ \mathrm{d}\theta \\
&= \int_0^\alpha a^3 \sin\theta[\pi(cos^2\theta - 1) + 2\pi\cos^2\theta]\ \mathrm{d}\theta \\
&= \int_0^\alpha a^3\pi(3\cos^3\theta - 1)\sin\theta\ \mathrm{d}\theta \\
&= \pi a^3[cos\theta - \cos^3\theta]_0^\alpha \\
&= \pi a^3 \cos\alpha \sin^2\alpha.
\end{aligned}$$

What happens when we change parametrization? Let $\mathbf{r}(u, v)$ and $\mathbf{r}(\tilde{u}, \tilde{v})$ be two regular parametrizations for the surface. By the chain rule,

$$\begin{aligned}
\frac{\partial\mathbf{r}}{\partial u} &= \frac{\partial\mathbf{r}}{\partial\tilde{u}}\frac{\partial\tilde{u}}{\partial u} + \frac{\partial\mathbf{r}}{\partial\tilde{v}}\frac{\partial\tilde{v}}{\partial u} \\
\frac{\partial\mathbf{r}}{\partial v} &= \frac{\partial\mathbf{r}}{\partial\tilde{u}}\frac{\partial\tilde{u}}{\partial v} + \frac{\partial\mathbf{r}}{\partial\tilde{v}}\frac{\partial\tilde{v}}{\partial v}
\end{aligned}$$

So

$$\frac{\partial\mathbf{r}}{\partial u} \times \frac{\partial\mathbf{r}}{\partial v} = \frac{\partial(\tilde{u}, \tilde{v})}{\partial(u, v)}\frac{\partial\mathbf{r}}{\partial\tilde{u}} \times \frac{\partial\mathbf{r}}{\partial\tilde{v}}$$

where $\frac{\partial(\tilde{u},\tilde{v})}{\partial(u,v)}$ is the Jacobian.

Since

$$\mathrm{d}\tilde{u}\,\mathrm{d}\tilde{v} = \frac{\partial(\tilde{u},\tilde{v})}{\partial(u,v)}\,\mathrm{d}u\,\mathrm{d}v,$$

We recover the formula

$$\mathrm{d}S = \left|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right| \mathrm{d}u\,\mathrm{d}v = \left|\frac{\partial \mathbf{r}}{\partial \tilde{u}} \times \frac{\partial \mathbf{r}}{\partial \tilde{v}}\right| \mathrm{d}\tilde{u}\,\mathrm{d}\tilde{v}.$$

Similarly, we have

$$\mathrm{d}\mathbf{S} = \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\,\mathrm{d}u\,\mathrm{d}v = \frac{\partial \mathbf{r}}{\partial \tilde{u}} \times \frac{\partial \mathbf{r}}{\partial \tilde{v}}\,\mathrm{d}\tilde{u}\,\mathrm{d}\tilde{v}.$$

provided (u, v) and $(\tilde{u}, \tilde{v})$ have the same orientation.

4.4 Change of variables in $\mathbb{R}^2$ and $\mathbb{R}^3$ revisited

In this section, we derive our change of variable formulae in a slightly different way.

Change of variable formula in $\mathbb{R}^2$

We first derive the 2D change of variable formula from the 3D surface integral formula.

Consider a subset S of the plane $\mathbb{R}^2$ parametrized by $\mathbf{r}(x(u,v), y(u,v))$. We can embed it to $\mathbb{R}^3$ as $\mathbf{r}(x(u,v), y(u,v), 0)$. Then

$$\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} = (0, 0, J),$$

with J being the Jacobian. Therefore

$$\int_S f(\mathbf{r})\,\mathrm{d}S = \int_D f(\mathbf{r}(u,v)) \left|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right| \mathrm{d}u\,\mathrm{d}v = \int_D f(\mathbf{r}(u,v))|J|\,\mathrm{d}u\,\mathrm{d}v,$$

and we recover the formula for changing variables in $\mathbb{R}^2$.

Change of variable formula in $\mathbb{R}^3$

In $\mathbb{R}^3$, suppose we have a volume parametrised by $\mathbf{r}(u, v, w)$. Then

$$\delta\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u}\delta u + \frac{\partial \mathbf{r}}{\partial v}\delta v + \frac{\partial \mathbf{r}}{\partial w}\delta w + o(\delta u, \delta v, \delta w).$$

Then the cuboid $\delta u, \delta v, \delta w$ in u, v, w space is mapped to a parallelopiped of volume

$$\delta V = \left|\frac{\partial \mathbf{r}}{\partial u}\delta u \cdot \left(\frac{\partial \mathbf{r}}{\partial v}\delta v \times \frac{\partial \mathbf{r}}{\partial w}\delta w\right)\right| = |J|\,\delta u\,\delta v\,\delta w.$$

So $\mathrm{d}V = |J|\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w$.

5 Geometry of curves and surfaces

Let $\mathbf{r}(s)$ be a curve parametrized by arclength s. Since $\mathbf{t}(s) = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}$ is a unit vector, $\mathbf{t} \cdot \mathbf{t} = 1$. Differentiating yields $\mathbf{t} \cdot \mathbf{t}' = 0$. So $\mathbf{t}'$ is a normal to the curve if $\mathbf{t}' \neq 0$.

We define the following:

Definition (Principal normal and curvature)**.** Write $\mathbf{t}' = \kappa \mathbf{n}$, where $\mathbf{n}$ is a unit vector and $\kappa > 0$. Then $\mathbf{n}(s)$ is called the *principal normal* and $\kappa(s)$ is called the *curvature*.

Note that we *must* be differentiating against s, not any other parametrization! If the curve is given in another parametrization, we can either change the parametrization or use the chain rule.

We take a curve that can Taylor expanded around $s = 0$. Then

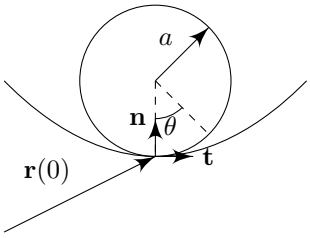
$$\mathbf{r}(s) = \mathbf{r}(0) + s\mathbf{r}'(0) + \frac{1}{2}s^2\mathbf{r}''(0) + O(s^3).$$

We know that $\mathbf{r}' = \mathbf{t}$ and $\mathbf{r}'' = \mathbf{t}'$. So we have

$$\mathbf{r}(s) = \mathbf{r}(0) + s\mathbf{t}(0) + \frac{1}{2}\kappa(0)s^2\mathbf{n} + O(s^3).$$

How can we interpret κ as the curvature? Suppose we want to approximate the curve near $\mathbf{r}(0)$ by a circle. We would expect a more "curved" curve would be approximated by a circle of smaller radius. So κ should be inversely proportional to the radius of the circle. In fact, we will show that $\kappa = 1/a$, where a is the radius of the best-fit circle.

Consider the vector equation for a circle passing through $\mathbf{r}(0)$ with radius a in the plane defined by $\mathbf{t}$ and $\mathbf{n}$.



Then the equation of the circle is

$$\mathbf{r} = \mathbf{r}(0) + a(1 - \cos\theta)\mathbf{n} + a\sin\theta\mathbf{t}.$$

We can expand this to obtain

$$\mathbf{r} = \mathbf{r}(0) + a\theta\mathbf{t} + \frac{1}{2}\theta^2 a\mathbf{n} + o(\theta^3).$$

Since the arclength $s = a\theta$, we obtain

$$\mathbf{r} = \mathbf{r}(0) + s\mathbf{t} + \frac{1}{2}\frac{1}{a}s^2\mathbf{n} + O(s^3).$$

As promised, $\kappa = 1/a$, for a the radius of the circle of best fit.

Definition (Radius of curvature)**.** The *radius of curvature* of a curve at a point $\mathbf{r}(s)$ is $1/\kappa(s)$.

Since we are in 3D, given $\mathbf{t}(s)$ and $\mathbf{n}(s)$, there is another normal to the curve. We can add a third normal to generate an orthonormal basis.

Definition (Binormal)**.** The *binormal* of a curve is $\mathbf{b} = \mathbf{t} \times \mathbf{n}$.

We can define the torsion similar to the curvature, but with the binormal instead of the tangent.[1]

Definition (Torsion)**.** Let $\mathbf{b}' = -\tau\mathbf{n}$. Then τ is the *torsion*.

Note that this makes sense, since $\mathbf{b}'$ is both perpendicular to $\mathbf{t}$ and $\mathbf{b}$, and hence must be in the same direction as $\mathbf{n}$. ($\mathbf{b}' = \mathbf{t}' \times \mathbf{n} + \mathbf{t} \times \mathbf{n}' = \mathbf{t} \times \mathbf{n}'$, so $\mathbf{b}'$ is perpendicular to $\mathbf{t}$; and $\mathbf{b} \cdot \mathbf{b} = 1 \Rightarrow \mathbf{b} \cdot \mathbf{b}' = 0$. So $\mathbf{b}'$ is perpendicular to $\mathbf{b}$).

The geometry of the curve is encoded in how this basis $(\mathbf{t}, \mathbf{n}, \mathbf{b})$ changes along it. This can be specified by two scalar functions of arc length — the curvature $\kappa(s)$ and the *torsion* $\tau(s)$ (which determines what the curve looks like to third order in its Taylor expansions and how the curve lifts out of the $\mathbf{t}, \mathbf{r}$ plane).

Surfaces and intrinsic geometry*

We can study the geometry of surfaces through curves which lie on them. At a given point P at a surface S with normal $\mathbf{n}$, consider a plane containing $\mathbf{n}$. The intersection of the plane with the surface yields a curve on the surface through P. This curve has a curvature κ at P.

If we choose different planes containing $\mathbf{n}$, we end up with different curves of different curvature. Then we define the following:

Definition (Principal curvature)**.** The *principal curvatures* of a surface at P are the minimum and maximum possible curvature of a curve through P, denoted $\kappa_{\min}$ and $\kappa_{\max}$ respectively.

Definition (Gaussian curvature)**.** The *Gaussian curvature* of a surface at a point P is $K = \kappa_{\min}\kappa_{\max}$.

Theorem (Theorema Egregium)**.** K is *intrinsic* to the surface S. It can be expressed in terms of lengths, angles etc. which are measured entirely on the surface. So K can be defined on an arbitrary surface without embedding it on a higher dimension surface.

The is the start of *intrinsic geometry*: if we embed a surface in Euclidean space, we can determine lengths, angles etc on it. But we don't have to do so — we can "live in " the surface and do geometry in it without an embedding.

For example, we can consider a geodesic triangle D on a surface S. It consists of three geodesics: shortest curves between two points.

Let θ_i be the interior angles of the triangle (defined by using scalar products of tangent vectors). Then

[1] This was not taught in lectures, but there is a question on the example sheet about the torsion, so I might as well include it here.

Theorem (Gauss-Bonnet theorem)**.**

$$\theta_1 + \theta_2 + \theta_3 = \pi + \int_D K \, \mathrm{d}A,$$

integrating over the area of the triangle.

6 Div, Grad, Curl and ∇

6.1 Div, Grad, Curl and ∇

Recalled that ∇f is given by $(\nabla f)_i = \frac{\partial f}{\partial x_i}$. We can regard this as obtained from the scalar field f by applying

$$\nabla = \mathbf{e}_i \frac{\partial}{\partial x_i}$$

for cartesian coordinates x_i and orthonormal basis $\mathbf{e}_i$, where $\mathbf{e}_i$ are orthonormal and right-handed, i.e. $\mathbf{e}_i \times \mathbf{e}_j = \varepsilon_{ijk}\mathbf{e}_k$ (it is left handed if $\mathbf{e}_i \times \mathbf{e}_j = -\varepsilon_{ijk}\mathbf{e}_k$).

We can alternatively write this as

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right).$$

∇ (*nabla* or *del*) is both an operator and a vector. We can apply it to a vector field $\mathbf{F}(\mathbf{r}) = \mathbf{F}_i(\mathbf{r})\mathbf{e}_i$ using the scalar or vector product.

Definition (Divergence)**.** The *divergence* or *div* of $\mathbf{F}$ is

$$\nabla \cdot \mathbf{F} = \frac{\partial F_i}{\partial x_i} = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3}.$$

Definition (Curl)**.** The *curl* of $\mathbf{F}$ is

$$\nabla \times \mathbf{F} = \varepsilon_{ijk}\frac{\partial F_k}{\partial x_j}\mathbf{e}_i = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix}$$

Example. Let $\mathbf{F} = (xe^z, y^2 \sin x, xyz)$. Then

$$\nabla \cdot \mathbf{F} = \frac{\partial}{\partial x}xe^z + \frac{\partial}{\partial y}y^2 \sin x + \frac{\partial}{\partial z}xyz = e^z + 2y\sin x + xy.$$

and

$$\begin{aligned} \nabla \times F &= \hat{\mathbf{i}}\left[\frac{\partial}{\partial y}(xyz) - \frac{\partial}{\partial z}(y^2 \sin x)\right] \\ &\quad + \hat{\mathbf{j}}\left[\frac{\partial}{\partial z}(xe^z) + \frac{\partial}{\partial x}(xyz)\right] \\ &\quad + \hat{\mathbf{k}}\left[\frac{\partial}{\partial x}(y^2 \sin x) - \frac{\partial}{\partial y}(xe^z)\right] \\ &= (xz, xe^z - yz, y^2 \cos x). \end{aligned}$$

Note that ∇ is an operator, so ordering is important. For example,

$$\mathbf{F} \cdot \nabla = F_i \frac{\partial}{\partial x_i}$$

is a scalar differential operator, and

$$\mathbf{F} \times \nabla = \mathbf{e}_k \varepsilon_{ijk} F_i \frac{\partial}{\partial x_j}$$

is a vector differential operator.

Proposition. Let f, g be scalar functions, $\mathbf{F}, \mathbf{G}$ be vector functions, and μ, λ be constants. Then

$$\begin{aligned}\nabla(\lambda f + \mu g) &= \lambda \nabla f + \mu \nabla g\\ \nabla \cdot (\lambda \mathbf{F} + \mu \mathbf{G}) &= \lambda \nabla \cdot \mathbf{F} + \mu \nabla \cdot \mathbf{G}\\ \nabla \times (\lambda \mathbf{F} + \mu \mathbf{G}) &= \lambda \nabla \times \mathbf{F} + \mu \nabla \times \mathbf{G}.\end{aligned}$$

Note that Grad and Div can be analogously defined in any dimension n, but curl is specific to $n = 3$ because it uses the vector product.

Example. Consider r^α with $r = |\mathbf{r}|$. We know that $\mathbf{r} = x_i \mathbf{e}_i$. So $r^2 = x_i x_i$. Therefore

$$2r\frac{\partial r}{\partial x_j} = 2x_j,$$

or

$$\frac{\partial r}{\partial x_i} = \frac{x_i}{r}.$$

So

$$\nabla r^\alpha = \mathbf{e}_i \frac{\partial}{\partial x_i}(r^\alpha) = \mathbf{e}_i \alpha r^{\alpha-1}\frac{\partial r}{\partial x_i} = \alpha r^{\alpha-2}\mathbf{r}.$$

Also,

$$\nabla \cdot \mathbf{r} = \frac{\partial x_i}{\partial x_i} = 3.$$

and

$$\nabla \times \mathbf{r} = \mathbf{e}_k \varepsilon_{ijk}\frac{\partial x_j}{\partial x_i} = 0.$$

Proposition. We have the following Leibnitz properties:

$$\begin{aligned}\nabla(fg) &= (\nabla f)g + f(\nabla g)\\ \nabla \cdot (f\mathbf{F}) &= (\nabla f)\cdot \mathbf{F} + f(\nabla \cdot \mathbf{F})\\ \nabla \times (f\mathbf{F}) &= (\nabla f) \times \mathbf{F} + f(\nabla \times \mathbf{F})\\ \nabla(\mathbf{F}\cdot\mathbf{G}) &= \mathbf{F}\times(\nabla\times\mathbf{G}) + \mathbf{G}\times(\nabla\times\mathbf{F}) + (\mathbf{F}\cdot\nabla)\mathbf{G} + (\mathbf{G}\cdot\nabla)\mathbf{F}\\ \nabla\times(\mathbf{F}\times\mathbf{G}) &= \mathbf{F}(\nabla\cdot\mathbf{G}) - \mathbf{G}(\nabla\cdot\mathbf{F}) + (\mathbf{G}\cdot\nabla)\mathbf{F} - (\mathbf{F}\cdot\nabla)\mathbf{G}\\ \nabla\cdot(\mathbf{F}\times\mathbf{G}) &= (\nabla\times\mathbf{F})\cdot\mathbf{G} - \mathbf{F}\cdot(\nabla\times\mathbf{G})\end{aligned}$$

which can be proven by brute-forcing with suffix notation and summation convention.

There is absolutely no point in memorizing these (at least the last three). They can be derived when needed via suffix notation.

Example.

$$\begin{aligned}\nabla\cdot(r^\alpha\mathbf{r}) &= (\nabla r^\alpha)\mathbf{r} + r^\alpha \nabla\cdot\mathbf{r}\\ &= (\alpha r^{\alpha-2}\mathbf{r})\cdot\mathbf{r} + r^\alpha(3)\\ &= (\alpha+3)r^\alpha\\ \nabla\times(r^\alpha\mathbf{r}) &= (\nabla(r^\alpha))\times\mathbf{r} + r^\alpha(\nabla\times\mathbf{r})\\ &= \alpha r^{\alpha-2}\mathbf{r}\times\mathbf{r}\\ &= \mathbf{0}\end{aligned}$$

6.2 Second-order derivatives

We have

Proposition.

$$\nabla \times (\nabla f) = 0$$
$$\nabla \cdot (\nabla \times \mathbf{F}) = 0$$

Proof. Expand out using suffix notation, noting that

$$\varepsilon_{ijk} \frac{\partial^2 f}{\partial x_i \partial x_j} = 0.$$

since if, say, $k = 3$, then

$$\varepsilon_{ijk} \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_1 \partial x_2} - \frac{\partial^2 f}{\partial x_2 \partial x_1} = 0.$$

□

The converse of each result holds for fields defined in all of $\mathbb{R}^3$:

Proposition. If $\mathbf{F}$ is defined in all of $\mathbb{R}^3$, then

$$\nabla \times \mathbf{F} = 0 \Rightarrow \mathbf{F} = \nabla f$$

for some f.

Definition (Conservative/irrotational field and scalar potential)**.** If $\mathbf{F} = \nabla f$, then f is the *scalar potential*. We say $\mathbf{F}$ is *conservative* or *irrotational*.

Similarly,

Proposition. If $\mathbf{H}$ is defined over all of $\mathbb{R}^3$ and $\nabla \cdot \mathbf{H} = 0$, then $\mathbf{H} = \nabla \times \mathbf{A}$ for some $\mathbf{A}$.

Definition (Solenoidal field and vector potential)**.** If $\mathbf{H} = \nabla \times \mathbf{A}$, $\mathbf{A}$ is the *vector potential* and $\mathbf{H}$ is said to be *solenoidal*.

Not that is is true only if $\mathbf{F}$ or $\mathbf{H}$ is defined on all of $\mathbb{R}^3$.

Definition (Laplacian operator)**.** The *Laplacian operator* is defined by

$$\nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial x_i \partial x_i} = \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^3} \right).$$

This operation is defined on both scalar and vector fields — on a scalar field,

$$\nabla^2 f = \nabla \cdot (\nabla f),$$

whereas on a vector field,

$$\nabla^2 \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla \times (\nabla \times \mathbf{A}).$$

7 Integral theorems

7.1 Statement and examples

There are three big integral theorems, known as Green's theorem, Stoke's theorem and Gauss' theorem. There are all generalizations of the fundamental theorem of calculus in some sense. In particular, they all say that an n dimensional integral of a derivative is equivalent to an $n-1$ dimensional integral of the original function.

We will first state all three theorems with some simple applications. In the next section, we will see that the three integral theorems are so closely related that it's easiest to show their equivalence first, and then prove just one of them.

7.1.1 Green's theorem (in the plane)

Theorem (Green's theorem)**.** For smooth functions $P(x,y)$, $Q(x,y)$ and A a bounded region in the (x,y) plane with boundary $\partial A = C$,

$$\int_A \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}A = \int_C (P\,\mathrm{d}x + Q\,\mathrm{d}y).$$

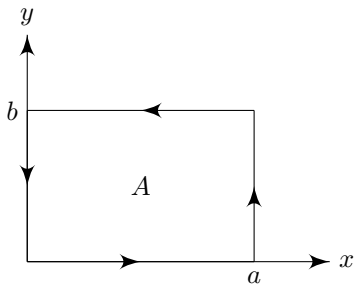
Here C is assumed to be piecewise smooth, non-intersecting closed curve, traversed anti-clockwise.

Example. Let $Q = xy^2$ and $P = x^2y$. If C is the parabola $y^2 = 4ax$ and the line $x = a$, both with $-2a \le y \le 2a$, then Green's theorem says

$$\int_A (y^2 - x^2)\,\mathrm{d}A = \int_C x^2\,\mathrm{d}x + xy^2\,\mathrm{d}y.$$

From example sheet 1, each side gives $\frac{104}{105}a^4$.

Example. Let A be a rectangle confined by $0 \le x \le a$ and $0 \le y \le b$.



Then Green's theorem follows directly from the fundamental theorem of calculus in 1D. We first consider the first term of Green's theorem:

$$\begin{aligned}\int -\frac{\partial P}{\partial y}\,\mathrm{d}A &= \int_0^a \int_0^b -\frac{\partial P}{\partial y}\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_0^a [-P(x,b) + P(x,0)]\,\mathrm{d}x \\ &= \int_C P\,\mathrm{d}x\end{aligned}$$

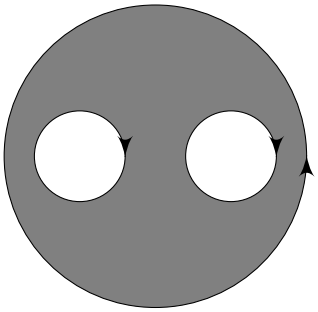
Note that we can convert the 1D integral in the second-to-last line to a line integral around the curve C, since the $P(x, 0)$ and $P(x, b)$ terms give the horizontal part of C, and the lack of $\mathrm{d}y$ term means that the integral is nil when integrating the vertical parts.

Similarly,

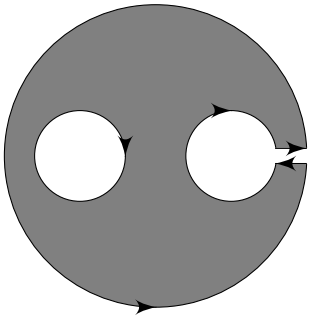
$$\int_A \frac{\partial Q}{\partial x}\,\mathrm{d}A = \int_C Q\,\mathrm{d}y.$$

Combining them gives Green's theorem.

Green's theorem also holds for a bounded region A, where the boundary ∂A consists of *disconnected* components (each piecewise smooth, non-intersecting and closed) with anti-clockwise orientation on the exterior, and clockwise on the interior boundary, e.g.



The orientation of the curve comes from imagining the surface as:



and take the limit as the gap shrinks to 0.

7.1.2 Stokes' theorem

Theorem (Stokes' theorem)**.** For a smooth vector field $\mathbf{F}(\mathbf{r})$,

$$\int_S \nabla \times \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \int_{\partial S} \mathbf{F} \cdot \mathrm{d}\mathbf{r},$$

where S is a smooth, bounded surface and ∂S is a piecewise smooth boundary of S.

The direction of the line integral is as follows: If we walk along C with $\mathbf{n}$ facing up, then the surface is on your left.

It also holds if ∂S is a collection of disconnected piecewise smooth closed curves, with the orientation determined in the same way as Green's theorem.

Example. Let S be the section of a sphere of radius a with $0 \leq \theta \leq \alpha$. In spherical coordinates,

$$\mathrm{d}\mathbf{S} = a^2 \sin\theta \mathbf{e}_r \, \mathrm{d}\theta \, \mathrm{d}\varphi.$$

Let $\mathbf{F} = (0, xz, 0)$. Then $\nabla \times \mathbf{F} = (-x, 0, z)$. We have previously shown that

$$\int_S \nabla \times \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \pi a^3 \cos\alpha \sin^2\alpha.$$

Our boundary ∂C is

$$\mathbf{r}(\varphi) = a(\sin\alpha\cos\varphi, \sin\alpha\sin\varphi, \cos\alpha).$$

The right hand side of Stokes' is

$$\begin{aligned}\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= \int_0^{2\pi} \underbrace{a\sin\alpha\cos\varphi}_{x} \underbrace{a\cos\alpha}_{z} \underbrace{a\sin\alpha\cos\varphi \, \mathrm{d}\varphi}_{\mathrm{d}y} \\ &= a^3 \sin^2\alpha\cos\alpha \int_0^{2\pi} \cos^2\varphi \, \mathrm{d}\varphi \\ &= \pi a^3 \sin^2\alpha\cos\alpha.\end{aligned}$$

So they agree.

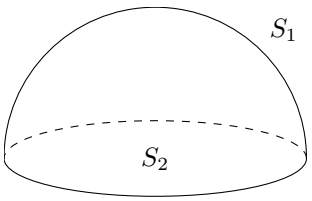
7.1.3 Divergence/Gauss theorem

Theorem (Divergence/Gauss theorem)**.** For a smooth vector field $\mathbf{F}(\mathbf{r})$,

$$\int_V \nabla \cdot \mathbf{F} \, \mathrm{d}V = \int_{\partial V} \mathbf{F} \cdot \mathrm{d}\mathbf{S},$$

where V is a bounded volume with boundary ∂V, a piecewise smooth, closed surface, with outward normal $\mathbf{n}$.

Example. Consider a hemisphere.



V is a solid hemisphere

$$x^2 + y^2 + z^2 \leq a^2, \quad z \geq 0,$$

and $\partial V = S_1 + S_2$, the hemisphere and the disc at the bottom.

Take $\mathbf{F} = (0, 0, z + a)$ and $\nabla \cdot \mathbf{F} = 1$. Then

$$\int_V \nabla \cdot \mathbf{F} \, \mathrm{d}V = \frac{2}{3}\pi a^3,$$

the volume of the hemisphere.

On S_1,

$$\mathrm{d}\mathbf{S} = \mathbf{n}\,\mathrm{d}S = \frac{1}{a}(x, y, z)\,\mathrm{d}S.$$

Then

$$\mathbf{F}\cdot\mathrm{d}\mathbf{S} = \frac{1}{a}z(z+a)\,\mathrm{d}S = \cos\theta a(\cos\theta+1)\underbrace{a^2\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi}_{\mathrm{d}S}.$$

Then

$$\begin{aligned}\int_{S_1}\mathbf{F}\cdot\mathrm{d}\mathbf{S} &= a^3\int_0^{2\pi}\mathrm{d}\varphi\int_0^{\pi/2}\sin\theta(\cos^2\theta+\cos\theta)\,\mathrm{d}\theta\\ &= 2\pi a^3\left[\frac{-1}{3}\cos^3\theta-\frac{1}{2}\cos^2\theta\right]_0^{\pi/2}\\ &= \frac{5}{3}\pi a^3.\end{aligned}$$

On S_2, $\mathrm{d}\mathbf{S} = \mathbf{n}\,\mathrm{d}S = -(0,0,1)\,\mathrm{d}S$. Then $\mathbf{F}\cdot\mathrm{d}\mathbf{S} = -a\,\mathrm{d}S$. So

$$\int_{S_2}\mathbf{F}\cdot\mathrm{d}\mathbf{S} = -\pi a^3.$$

So

$$\int_{S_1}\mathbf{F}\cdot\mathrm{d}\mathbf{S}+\int_{S_2}\mathbf{F}\cdot\mathrm{d}\mathbf{S} = \left(\frac{5}{3}-1\right)\pi a^3 = \frac{2}{3}\pi a^3,$$

in accordance with Gauss' theorem.

7.2 Relating and proving integral theorems

We will first show the following two equivalences:

- Stokes' theorem $\Leftrightarrow$ Green's theorem
- 2D divergence theorem $\Leftrightarrow$ Greens' theorem

Then we prove the 2D version of divergence theorem directly to show that all of the above hold. A sketch of the proof of the 3D version of divergence theorem will be provided, because it is simply a generalization of the 2D version, except that the extra dimension makes the notation tedious and difficult to follow.

Proposition. Stokes' theorem $\Rightarrow$ Green's theorem

Proof. Stokes' theorem talks about 3D surfaces and Green's theorem is about 2D regions. So given a region A on the (x, y) plane, we pretend that there is a third dimension and apply Stokes' theorem to derive Green's theorem.

Let A be a region in the (x, y) plane with boundary $C = \partial A$, parametrised by arc length, $(x(s), y(s), 0)$. Then the tangent to C is

$$\mathbf{t} = \left(\frac{\mathrm{d}x}{\mathrm{d}s}, \frac{\mathrm{d}y}{\mathrm{d}s}, 0\right).$$

Given any $P(x, y)$ and $Q(x, y)$, we can consider the vector field

$$\mathbf{F} = (P, Q, 0),$$

So

$$\nabla \times \mathbf{F} = \left(0, 0, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right).$$

Then the left hand side of Stokes is

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_C \mathbf{F} \cdot \mathbf{t}\, \mathrm{d}s = \int_C P\, \mathrm{d}x + Q\, \mathrm{d}y,$$

and the right hand side is

$$\int_A (\nabla \times \mathbf{F}) \cdot \hat{\mathbf{k}}\, \mathrm{d}A = \int_A \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}A. \qquad \square$$

Proposition. Green's theorem $\Rightarrow$ Stokes' theorem.

Proof. Green's theorem describes a 2D region, while Stokes' theorem describes a 3D surface $\mathbf{r}(u, v)$. Hence to use Green's to derive Stokes' we need find some 2D thing to act on. The natural choice is the parameter space, u, v.

Consider a parametrised surface $S = \mathbf{r}(u, v)$ corresponding to the region A in the u, v plane. Write the boundary as $\partial A = (u(t), v(t))$. Then $\partial S = \mathbf{r}(u(t), v(t))$.

We want to prove

$$\int_{\partial S} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_S (\nabla \times \mathbf{F}) \cdot \mathrm{d}\mathbf{S}$$

given

$$\int_{\partial A} F_u\, \mathrm{d}u + F_v\, \mathrm{d}v = \int_A \left(\frac{\partial F_v}{\partial u} - \frac{\partial F_u}{\partial v}\right) \mathrm{d}A.$$

Doing some pattern-matching, we want

$$\mathbf{F} \cdot \mathrm{d}\mathbf{r} = F_u\, \mathrm{d}u + F_v\, \mathrm{d}v$$

for some F_u and F_v.

By the chain rule, we know that

$$\mathrm{d}\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u}\mathrm{d}u + \frac{\partial \mathbf{r}}{\partial v}\mathrm{d}v.$$

So we choose

$$F_u = \mathbf{F} \cdot \frac{\partial \mathbf{r}}{\partial u}, \quad F_v = \mathbf{F} \cdot \frac{\partial \mathbf{r}}{\partial v}.$$

This choice matches the left hand sides of the two equations.

To match the right, recall that

$$(\nabla \times \mathbf{F}) \cdot \mathrm{d}\mathbf{S} = (\nabla \times \mathbf{F}) \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right) \mathrm{d}u\, \mathrm{d}v.$$

Therefore, for the right hand sides to match, we want

$$\frac{\partial F_v}{\partial u} - \frac{\partial F_u}{\partial v} = (\nabla \times \mathbf{F}) \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right). \qquad (*)$$

Fortunately, this is true. Unfortunately, the proof involves complicated suffix notation and summation convention:

$$\frac{\partial F_v}{\partial u} = \frac{\partial}{\partial u}\left(\mathbf{F} \cdot \frac{\partial \mathbf{r}}{\partial v}\right) = \frac{\partial}{\partial u}\left(F_i \frac{\partial x_i}{\partial v}\right) = \left(\frac{\partial F_i}{\partial x_j}\frac{\partial x_j}{\partial u}\right)\frac{\partial x_i}{\partial v} + F_i \frac{\partial x_i}{\partial u \partial v}.$$

Similarly,

$$\frac{\partial F_u}{\partial u} = \frac{\partial}{\partial u}\left(\mathbf{F}\cdot\frac{\partial \mathbf{r}}{\partial u}\right) = \frac{\partial}{\partial u}\left(F_j\frac{\partial x_j}{\partial u}\right) = \left(\frac{\partial F_j}{\partial x_i}\frac{\partial x_i}{\partial v}\right)\frac{\partial x_j}{\partial u} + F_i\frac{\partial x_i}{\partial u\partial v}.$$

So

$$\frac{\partial F_v}{\partial u} - \frac{\partial F_u}{\partial v} = \frac{\partial x_j}{\partial u}\frac{\partial x_i}{\partial v}\left(\frac{\partial F_i}{\partial x_j} - \frac{\partial F_j}{\partial x_i}\right).$$

This is the left hand side of $(*)$.

The right hand side of $(*)$ is

$$\begin{aligned}(\nabla\times\mathbf{F})\cdot\left(\frac{\partial \mathbf{r}}{\partial u}\times\frac{\partial \mathbf{r}}{\partial v}\right) &= \varepsilon_{ijk}\frac{\partial F_j}{\partial x_i}\varepsilon_{kpq}\frac{\partial x_p}{\partial u}\frac{\partial x_q}{\partial v}\\ &= (\delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp})\frac{\partial F_j}{\partial x_i}\frac{\partial x_p}{\partial u}\frac{\partial x_q}{\partial v}\\ &= \left(\frac{\partial F_j}{\partial x_i} - \frac{\partial F_i}{\partial x_j}\right)\frac{\partial x_i}{\partial u}\frac{\partial x_j}{\partial v}.\end{aligned}$$

So they match. Therefore, given our choice of F_u and F_v, Green's theorem translates to Stokes' theorem. □

Proposition. Greens theorem $\Leftrightarrow$ 2D divergence theorem.

Proof. The 2D divergence theorem states that

$$\int_A (\nabla\cdot\mathbf{G})\,\mathrm{d}A = \int_{\partial A}\mathbf{G}\cdot\mathbf{n}\,\mathrm{d}s.$$

with an outward normal $\mathbf{n}$.

Write $\mathbf{G}$ as $(Q, -P)$. Then

$$\nabla\cdot\mathbf{G} = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}.$$

Around the curve $\mathbf{r}(s) = (x(s), y(s))$, $\mathbf{t}(s) = (x'(s), y'(s))$. Then the normal, being tangent to $\mathbf{t}$, is $\mathbf{n}(s) = (y'(s), -x'(s))$ (check that it points outwards!). So

$$\mathbf{G}\cdot\mathbf{n} = P\frac{\mathrm{d}x}{\mathrm{d}s} + Q\frac{\mathrm{d}y}{\mathrm{d}s}.$$

Then we can expand out the integrals to obtain

$$\int_C \mathbf{G}\cdot\mathbf{n}\,\mathrm{d}s = \int_C P\,\mathrm{d}x + Q\,\mathrm{d}y,$$

and

$$\int_A(\nabla\cdot\mathbf{G})\,\mathrm{d}A = \int_A\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\,\mathrm{d}A.$$

Now 2D version of Gauss' theorem says the two LHS are the equal, and Green's theorem says the two RHS are equal. So the result follows. □

Proposition. 2D divergence theorem.

$$\int_A(\nabla\cdot\mathbf{G})\,\mathrm{d}A = \int_{C=\partial A}\mathbf{G}\cdot\mathbf{n}\,\mathrm{d}s.$$

Proof. For the sake of simplicity, we assume that $\mathbf{G}$ only has a vertical component, noting that the same proof works for purely horizontal $\mathbf{G}$, and an arbitrary $\mathbf{G}$ is just a linear combination of the two.

Furthermore, we assume that A is a simple, convex shape. A more complicated shape can be cut into smaller simple regions, and we can apply the simple case to each of the small regions.

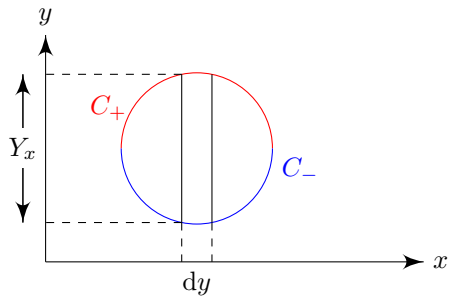
Suppose $\mathbf{G} = G(x, y)\hat{\mathbf{j}}$. Then

$$\nabla \cdot \mathbf{G} = \frac{\partial G}{\partial y}.$$

Then

$$\int_A \nabla \cdot \mathbf{G} \, \mathrm{d}A = \int_X \left(\int_{Y_x} \frac{\partial G}{\partial y} \, \mathrm{d}y \right) \, \mathrm{d}x.$$

Now we divide A into an upper and lower part, with boundaries $C_+ = y_+(x)$ and $C_- = y_-(x)$ respectively. Since I cannot draw, A will be pictured as a circle, but the proof is valid for any simple convex shape.



We see that the boundary of Y_x at any specific x is given by $y_-(x)$ and $y_+(x)$. Hence by the Fundamental theorem of Calculus,

$$\int_{Y_x} \frac{\partial G}{\partial y} \, \mathrm{d}y = \int_{y_-(x)}^{y_+(x)} \frac{\partial G}{\partial y} \, \mathrm{d}y = G(x, y_+(x)) - G(x, y_-(x)).$$

To compute the full area integral, we want to integrate over all x. However, the divergence theorem talks in terms of $\mathrm{d}s$, not $\mathrm{d}x$. So we need to find some way to relate $\mathrm{d}s$ and $\mathrm{d}x$. If we move a distance δs, the change in x is $\delta s \cos\theta$, where θ is the angle between the tangent and the horizontal. But θ is also the angle between the normal and the vertical. So $\cos\theta = \mathbf{n} \cdot \hat{\mathbf{j}}$. Therefore $\mathrm{d}x = \hat{\mathbf{j}} \cdot \mathbf{n} \, \mathrm{d}s$.

In particular, $G \, \mathrm{d}x = G\hat{\mathbf{j}} \cdot \mathbf{n} \, \mathrm{d}s = \mathbf{G} \cdot \mathbf{n} \, \mathrm{d}s$, since $\mathbf{G} = G\hat{\mathbf{j}}$.

However, at C_-, $\mathbf{n}$ points downwards, so $\mathbf{n} \cdot \hat{\mathbf{j}}$ happens to be negative. So, actually, at C_-, $\mathrm{d}x = -\mathbf{G} \cdot \mathbf{n} \, \mathrm{d}s$.

Therefore, our full integral is

$$\begin{aligned}\int_A \nabla \cdot \mathbf{G}\,\mathrm{d}A &= \int_X \left(\int_{y_x} \frac{\partial G}{\partial y}\,\mathrm{d}Y\right)\,\mathrm{d}x\\ &= \int_X G(x, y_+(x)) - G(x, y_-(x))\,\mathrm{d}x\\ &= \int_{C_+} \mathbf{G}\cdot\mathbf{n}\,\mathrm{d}s + \int_{C_-} \mathbf{G}\cdot\mathbf{n}\,\mathrm{d}s\\ &= \int_C \mathbf{G}\cdot\mathbf{n}\,\mathrm{d}s. \qquad \square\end{aligned}$$

To prove the 3D version, we again consider $\mathbf{F} = F(x, y, z)\hat{\mathbf{k}}$, a purely vertical vector field. Then

$$\int_V \nabla\cdot\mathbf{F}\,\mathrm{d}V = \int_D \left(\int_{Z_{xy}} \frac{\partial F}{\partial z}\,\mathrm{d}z\right)\,\mathrm{d}A.$$

Again, split $S = \partial V$ into the top and bottom parts S_+ and S_- (ie the parts with $\hat{\mathbf{k}}\cdot\mathbf{n} \geq 0$ and $\hat{\mathbf{k}}\cdot\mathbf{n} < 0$), and parametrize by $z_+(x, y)$ and $z_-(x, y)$. Then the integral becomes

$$\int_V \nabla\cdot\mathbf{F}\,\mathrm{d}V = \int_D (F(x, y, z_+) - F(x, y, z_-))\,\mathrm{d}A = \int_S \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}S.$$

8 Some applications of integral theorems

8.1 Integral expressions for div and curl

We can use these theorems to come up with alternative definitions of the div and curl. The advantage of these alternative definitions is that they do not require a choice of coordinate axes. They also better describe how we should interpret div and curl.

Gauss' theorem for $\mathbf{F}$ in a small volume V containing $\mathbf{r}_0$ gives

$$\int_{\partial V} \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \int_V \nabla \cdot \mathbf{F} \, \mathrm{d}V \approx (\nabla \cdot \mathbf{F})(\mathbf{r}_0) \, \mathrm{vol}(V).$$

We take the limit as $V \to 0$ to obtain

Proposition.

$$(\nabla \cdot \mathbf{F})(\mathbf{r}_0) = \lim_{\mathrm{diam}(V) \to 1} \frac{1}{\mathrm{vol}(V)} \int_{\partial V} \mathbf{F} \cdot \mathrm{d}\mathbf{S},$$

where the limit is taken over volumes containing the point $\mathbf{r}_0$.

Similarly, Stokes' theorem gives, for A a surface containing the point $\mathbf{r}_0$,

$$\int_{\partial A} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_A (\nabla \times \mathbf{F}) \cdot \mathbf{n} \, \mathrm{d}A \approx \mathbf{n} \cdot (\nabla \times \mathbf{F})(\mathbf{r}_0) \, \mathrm{area}(A).$$

So

Proposition.

$$\mathbf{n} \cdot (\nabla \times \mathbf{F})(\mathbf{r}_0) = \lim_{\mathrm{diam}(A) \to 0} \frac{1}{\mathrm{area}(A)} \int_{\partial A} \mathbf{F} \cdot \mathrm{d}\mathbf{r},$$

where the limit is taken over all surfaces A containing $\mathbf{r}_0$ with normal $\mathbf{n}$.

These are coordinate-independent definitions of div and curl.

Example. Suppose $\mathbf{u}$ is a velocity field of fluid flow. Then

$$\int_S \mathbf{u} \cdot \mathrm{d}\mathbf{S}$$

is the rate of which fluid crosses S. Taking V to be the volume occupied by a fixed quantity of fluid material, we have

$$\dot{V} = \int_{\partial V} \mathbf{u} \cdot \mathrm{d}\mathbf{S}$$

Then, at $\mathbf{r}_0$,

$$\nabla \cdot \mathbf{u} = \lim_{V \to 0} \frac{\dot{V}}{V},$$

the relative rate of change of volume. For example, if $\mathbf{u}(\mathbf{r}) = \alpha \mathbf{r}$ (ie fluid flowing out of origin), then $\nabla \cdot \mathbf{u} = 3\alpha$, which increases at a constant rate everywhere.

Alternatively, take a planar area A to be a disc of radius a. Then

$$\int_{\partial A} \mathbf{u} \cdot \mathrm{d}\mathbf{r} = \int_{\partial A} \mathbf{u} \cdot \mathbf{t} \, \mathrm{d}s = 2\pi a \times \text{average of } \mathbf{u} \cdot \mathbf{t} \text{ around the circumference.}$$

($\mathbf{u} \cdot \mathbf{t}$ is the component of $\mathbf{u}$ which is tangential to the boundary) We define the quantity

$$\omega = \frac{1}{a} \times (\text{average of } \mathbf{u} \cdot \mathbf{t}).$$

This is the local angular velocity of the current. As $a \to 0$, $\frac{1}{a} \to \infty$, but the average of $\mathbf{u} \cdot \mathbf{t}$ will also decrease since a smooth field is less "twirly" if you look closer. So ω tends to some finite value as $a \to 0$. We have

$$\int_{\partial A} \mathbf{u} \cdot \mathrm{d}\mathbf{r} = 2\pi a^2 \omega.$$

Recall that

$$\mathbf{n} \cdot \nabla \times \mathbf{u} = \lim_{A \to 0} \frac{1}{\pi a^2} \int_{\partial A} \mathbf{u} \cdot \mathrm{d}\mathbf{r} = 2\omega,$$

ie twice the local angular velocity. For example, if you have a washing machine rotating at a rate of $\boldsymbol{\omega}$, Then the velocity $\mathbf{u} = \boldsymbol{\omega} \times \mathbf{r}$. Then the curl is

$$\nabla \times (\boldsymbol{\omega} \times \mathbf{r}) = 2\boldsymbol{\omega},$$

which is twice the angular velocity.

8.2 Conservative fields and scalar products

Definition (Conservative field)**.** A vector field $\mathbf{F}$ is *conservative* if

(i) $\mathbf{F} = \nabla f$ for some scalar field f; or

(ii) $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ is independent of C, for fixed end points and orientation; or

(iii) $\nabla \times \mathbf{F} = 0$.

In $\mathbb{R}^3$, all three formulations are equivalent.

We have previously shown (i) $\Rightarrow$ (ii) since

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = f(\mathbf{b}) - f(\mathbf{a}).$$

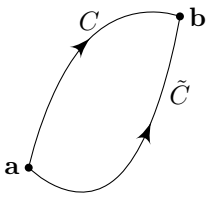
We have also shown that (i) $\Rightarrow$ (iii) since

$$\nabla \times (\nabla f) = 0.$$

So we want to show that (iii) $\Rightarrow$ (ii) and (ii) $\Rightarrow$ (i)

Proposition. If (iii) $\nabla \times \mathbf{F} = 0$, then (ii) $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ is independent of C.

Proof. Given $\mathbf{F}(\mathbf{r})$ satisfying $\nabla \times \mathbf{F} = 0$, let C and $\tilde{C}$ be any two curves from $\mathbf{a}$ to $\mathbf{b}$.



If S is any surface with boundary $\partial S = C - \tilde{C}$, By Stokes' theorem,

$$\int_S \nabla \times \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \int_{\partial S} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} - \int_{\tilde{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}.$$

But $\nabla \times \mathbf{F} = 0$. So

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} - \int_{\tilde{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0,$$

or

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_{\tilde{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}. \qquad \square$$

Proposition. If (ii) $\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ is independent of C for fixed end points and orientation, then (i) $\mathbf{F} = \nabla f$ for some scalar field f.

Proof. We fix $\mathbf{a}$ and define $f(\mathbf{r}) = \int_C \mathbf{F}(\mathbf{r}') \cdot \mathrm{d}\mathbf{r}'$ for any curve from $\mathbf{a}$ to $\mathbf{r}$. Assuming (ii), f is well-defined. For small changes $\mathbf{r}$ to $\mathbf{r} + \delta\mathbf{r}$, there is a small extension of C by δC. Then

$$\begin{aligned} f(\mathbf{r} + \delta\mathbf{r}) &= \int_{C+\delta C} \mathbf{F}(\mathbf{r}') \cdot \mathrm{d}\mathbf{r}' \\ &= \int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}' + \int_{\delta C} \mathbf{F} \cdot \mathrm{d}\mathbf{r}' \\ &= f(\mathbf{r}) + \mathbf{F}(\mathbf{r}) \cdot \delta\mathbf{r} + o(\delta\mathbf{r}). \end{aligned}$$

So

$$\delta f = f(\mathbf{r} + \delta\mathbf{r}) - f(\mathbf{r}) = \mathbf{F}(\mathbf{r}) \cdot \delta\mathbf{r} + o(\delta\mathbf{r}).$$

But the definition of grad is exactly

$$\delta f = \nabla f \cdot \delta\mathbf{r} + o(\delta\mathbf{r}).$$

So we have $\mathbf{F} = \nabla f$. $\square$

Note that these results assume $\mathbf{F}$ is defined on the whole of $\mathbb{R}^3$. It also works of $\mathbf{F}$ is defined on a *simply connected* domain D, ie a subspace of $\mathbb{R}^3$ without holes. By definition, this means that any two curves $C, \tilde{C}$ with fixed end points can be smoothly deformed into one another (alternatively, any loop can be shrunk into a point).

If we have a smooth transformation from C to $\tilde{C}$, the process sweeps out a surface bounded by C and $\tilde{C}$. This is required by the proof that (iii) $\Rightarrow$ (ii).

If D is not simply connected, then we obtain a multi-valued $f(\mathbf{r})$ on D in general (for the proof (ii) $\Rightarrow$ (i)). However, we can choose to restrict to a subset $D_0 \subseteq D$ such that $f(\mathbf{r})$ is single-valued on D_0.

Example. Take

$$\mathbf{F} = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}, 0\right).$$

This obeys $\nabla \times \mathbf{F} = 0$, and is defined on $D = \mathbb{R}^3 \setminus \{z\text{-axis}\}$, which is not simply-connected. We can also write

$$\mathbf{F} = \nabla f,$$

where

$$f = \tan^{-1} \frac{y}{x}.$$

which is multi-valued. If we integrate it about the closed loop $x^2 + y^2 = 1, z = 0$, i.e. a circle about the z axis, the integral gives 2π, as opposed to the expected 0 for a conservative force. This shows that the simply-connected-domain criterion is important!

However f can be single-valued if we restrict it to

$$D_0 = \mathbb{R}^3 - \{\text{half-plane } x \geq 0,\ y = 0\},$$

which is simply-connected. (Draw and check!) Any closed curve we can draw in this area will have an integral of 0 (the circle mentioned above will no longer be closed!).

8.3 Conservation laws

Definition (Conservation equation)**.** Suppose we are interested in a quantity Q. Let $\rho(\mathbf{r}, t)$ be the amount of stuff per unit volume and $\mathbf{j}(\mathbf{r}, t)$ be the flow rate of the quantity (eg if Q is charge, j is the current density).

The conservation equation is

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0.$$

This is stronger than the claim that the total amount of Q in the universe is fixed. It says that Q cannot just disappear here and appear elsewhere. It must continuously flow out.

In particular, let V be a fixed time-independent volume with boundary $S = \partial V$. Then

$$Q(t) = \int_V \rho(\mathbf{r}, t)\, \mathrm{d}V$$

Then the rate of change of amount of Q in V is

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \int_V \frac{\partial \rho}{\partial t}\, \mathrm{d}V = -\int_V \nabla \cdot \mathbf{j}\, \mathrm{d}V = -\int_S \mathbf{j} \cdot \mathrm{d}\mathbf{s}.$$

by divergence theorem. So this states that the rate of change of the quantity Q in V is the flux of the stuff flowing out of the surface. ie Q cannot just disappear but must smoothly flow out.

In particular, if V is the whole universe (ie $\mathbb{R}^3$), and $\mathbf{j} \to 0$ sufficiently rapidly as $|\mathbf{r}| \to \infty$, then we calculate the total amount of Q in the universe by taking V to be a solid sphere of radius R, and take the limit as $R \to \infty$. Then the surface integral $\to 0$, and the equation states that

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = 0,$$

Example. If $\rho(\mathbf{r}, t)$ is the charge density (i.e. $\rho\delta V$ is the amount of charge in a small volume δV), then $Q(t)$ is the total charge in V. $\mathbf{j}(\mathbf{r}, t)$ is the electric current density. So $\mathbf{j} \cdot \mathrm{d}\mathbf{S}$ is the charge flowing through δS per unit time.

Example. Let $\mathbf{j} = \rho\mathbf{u}$ with $\mathbf{u}$ being the velocity field. Then $(\rho\mathbf{u}\ \delta t) \cdot \delta\mathbf{S}$ is equal to the mass of fluid crossing δS in time δt. So

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = -\int_S \mathbf{j} \cdot \mathrm{d}\mathbf{S}$$

does indeed imply the conservation of mass. The conservation equation in this case is

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho\mathbf{u}) = 0$$

For the case where ρ is constant and uniform (i.e. independent of $\mathbf{r}$ and t), we get that $\nabla \cdot \mathbf{u} = 0$. We say that the fluid is *incompressible*.

9 Orthogonal curvilinear coordinates

9.1 Line, area and volume elements

In this chapter, we study funny coordinate systems. A coordinate system is, roughly speaking, a way to specify a point in space by a set of (usually 3) numbers. We can think of this as a function $\mathbf{r}(u, v, w)$.

By the chain rule, we have

$$\mathrm{d}\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u}\mathrm{d}u + \frac{\partial \mathbf{r}}{\partial v}\mathrm{d}v + \frac{\partial \mathbf{r}}{\partial w}\mathrm{d}w$$

For a good parametrization,

$$\frac{\partial \mathbf{r}}{\partial u} \cdot \left(\frac{\partial \mathbf{r}}{\partial v} \times \frac{\partial \mathbf{r}}{\partial w}\right) \neq 0,$$

i.e. $\frac{\partial \mathbf{r}}{\partial u}, \frac{\partial \mathbf{r}}{\partial v}$ and $\frac{\partial \mathbf{r}}{\partial w}$ are linearly independent. These vectors are tangent to the curves parametrized by u, v, w respectively when the other two are being fixed.

Even better, they should be orthogonal:

Definition (Orthogonal curvilinear coordinates)**.** u, v, w are *orthogonal curvilinear* if the tangent vectors are orthogonal.

We can then set

$$\frac{\partial \mathbf{r}}{\partial u} = h_u \mathbf{e}_u, \quad \frac{\partial \mathbf{r}}{\partial v} = h_v \mathbf{e}_v, \quad \frac{\partial \mathbf{r}}{\partial w} = h_w \mathbf{e}_w,$$

with $h_u, h_v, h_w > 0$ and $\mathbf{e}_u, \mathbf{e}_v, \mathbf{e}_w$ form an *orthonormal* right-handed basis (i.e. $\mathbf{e}_u \times \mathbf{e}_v = \mathbf{e}_w$). Then

$$\mathrm{d}\mathbf{r} = h_u \mathbf{e}_u \, \mathrm{d}u + h_v \mathbf{e}_v \, \mathrm{d}v + h_w \mathbf{e}_w \, \mathrm{d}w,$$

and h_u, h_v, h_w determine the changes in length along each orthogonal direction resulting from changes in u, v, w. Note that clearly by definition, we have

$$h_u = \left|\frac{\partial \mathbf{r}}{\partial u}\right|.$$

Example.

(i) In cartesian coordinates, $\mathbf{r}(x, y, z) = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$. Then $h_x = h_y = h_z = 1$, and $\mathbf{e}_x = \hat{\mathbf{i}}, \mathbf{e}_y = \hat{\mathbf{j}}$ and $\mathbf{e}_z = \hat{\mathbf{k}}$.

(ii) In cylindrical polars, $\mathbf{r}(\rho, \varphi, z) = \rho[\cos\varphi\hat{\mathbf{i}} + \sin\varphi\hat{\mathbf{j}}] + z\hat{\mathbf{k}}$. Then $h_\rho = h_z = 1$, and

$$h_\varphi = \left|\frac{\partial \mathbf{r}}{\partial \varphi}\right| = |(-\rho\sin\varphi, \rho\sin\varphi, 0)| = \rho.$$

The basis vectors $\mathbf{e}_\rho, \mathbf{e}_\varphi, \mathbf{e}_z$ are as in section 1.

(iii) In spherical polars,

$$\mathbf{r}(r, \theta, \varphi) = r(\cos\varphi\sin\theta\hat{\mathbf{i}} + \sin\theta\sin\varphi\hat{\mathbf{j}} + \cos\theta\hat{\mathbf{k}}).$$

Then $h_r = 1, h_\theta = r$ and $h_\varphi = r\sin\theta$.

Consider a surface with w constant and parametrised by u and v. The vector area element is

$$\mathrm{d}\mathbf{S} = \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\,\mathrm{d}u\,\mathrm{d}v = h_u\mathbf{e}_u \times h_v\mathbf{e}_v\,\mathrm{d}u\,\mathrm{d}v = h_uh_v\mathbf{e}_w\,\mathrm{d}u\,\mathrm{d}v.$$

We interpret this as δS having a small rectangle with sides approximately $h_u\delta u$ and $h_v\delta v$. The volume element is

$$\mathrm{d}V = \frac{\partial \mathbf{r}}{\partial u} \cdot \left(\frac{\partial \mathbf{r}}{\partial v} \times \frac{\partial \mathbf{r}}{\partial w}\right)\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w = h_uh_vh_w\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w,$$

i.e. a small cuboid with sides $h_u\delta_u$, $h_v\delta_v$ and $h_w\delta_w$ respectively.

9.2 Grad, Div and Curl

Consider $f(\mathbf{r}(u,v,w))$ and compare

$$\mathrm{d}f = \frac{\partial f}{\partial u}\,\mathrm{d}u + \frac{\partial f}{\partial v}\,\mathrm{d}v + \frac{\partial f}{\partial w}\,\mathrm{d}w,$$

with $\mathrm{d}f = (\nabla f)\cdot\mathrm{d}\mathbf{r}$. Since we know that

$$\mathrm{d}\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u}\,\mathrm{d}u + \frac{\partial \mathbf{r}}{\partial v}\,\mathrm{d}v + \frac{\partial \mathbf{r}}{\partial w}\,\mathrm{d}w = h_u\mathbf{e}_u\,\mathrm{d}u + h_v\mathbf{e}_v\,\mathrm{d}v + h_w\mathbf{e}_w\,\mathrm{d}v,$$

we can compare the terms to know that

Proposition.

$$\nabla f = \frac{1}{h_u}\frac{\partial f}{\partial u}\mathbf{e}_u + \frac{1}{h_v}\frac{\partial f}{\partial v}\mathbf{e}_v + \frac{1}{h_w}\frac{\partial f}{\partial w}\mathbf{e}_w.$$

Example. Take $f = r\sin\theta\cos\varphi$ in spherical polars. Then

$$\begin{aligned}\nabla f &= \sin\theta\cos\varphi\,\mathbf{e}_r + \frac{1}{r}(r\cos\theta\cos\varphi)\,\mathbf{e}_\theta + \frac{1}{r\sin\theta}(-r\sin\theta\sin\varphi)\,\mathbf{e}_\varphi\\ &= \cos\varphi(\sin\theta\,\mathbf{e}_r + \cos\theta\,\mathbf{e}_\theta) - \sin\varphi\,\mathbf{e}_\varphi.\end{aligned}$$

Then we know that the differential operator is

Proposition.

$$\nabla = \frac{1}{h_u}\mathbf{e}_u\frac{\partial}{\partial u} + \frac{1}{h_v}\mathbf{e}_v\frac{\partial}{\partial v} + \frac{1}{h_w}\mathbf{e}_w\frac{\partial}{\partial w}.$$

We can apply this to a vector field

$$\mathbf{F} = F_u\mathbf{e}_u + F_v\mathbf{e}_v + F_w\mathbf{e}_w$$

using scalar or vector products to obtain

Proposition.

$$\begin{aligned}\nabla\times\mathbf{F} &= \frac{1}{h_vh_w}\left[\frac{\partial}{\partial v}(h_wF_w) - \frac{\partial}{\partial w}(h_vF_v)\right]\mathbf{e}_u + \text{ two similar terms}\\ &= \frac{1}{h_uh_vh_w}\begin{vmatrix} h_u\mathbf{e}_u & h_v\mathbf{e}_v & h_w\mathbf{e}_w\\ \frac{\partial}{\partial u} & \frac{\partial}{\partial v} & \frac{\partial}{\partial w}\\ h_uF_u & h_vF_v & h_wF_w\end{vmatrix}\end{aligned}$$

and

$$\nabla\cdot\mathbf{F} = \frac{1}{h_uh_vh_w}\left[\frac{\partial}{\partial u}(h_vh_wF_u) + \text{ two similar terms}\right].$$

There are several ways to obtain these formulae. We can

Proof. (non-examinable)

(i) Apply $\nabla\cdot$ or $\nabla\times$ and differentiate the basis vectors explicitly.

(ii) First, apply $\nabla\cdot$ or $\nabla\times$, but calculate the result by writing $\mathbf{F}$ in terms of $\nabla u, \nabla v$ and ∇w in a suitable way. Then use $\nabla\times\nabla f = 0$ and $\nabla\cdot(\nabla\times f) = 0$.

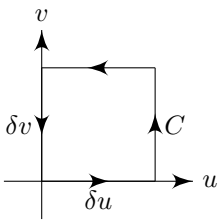
(iii) Use the integral expressions for div and curl.

Recall that

$$\mathbf{n}\cdot\nabla\times\mathbf{F} = \lim_{A\to 0}\frac{1}{A}\int_{\partial A}\mathbf{F}\cdot\mathrm{d}\mathbf{r}.$$

So to calculate the curl, we first find the $\mathbf{e}_w$ component.

Consider an area with W fixed and change u by δu and v by δv. Then this has an area of $h_u h_v \delta u \delta v$ with normal $\mathbf{e}_w$. Let C be its boundary.



We then integrate around the curve C. We split the curve C up into 4 parts (corresponding to the four sides), and take linear approximations by assuming F and h are constant when moving through each horizontal/vertical segment.

$$\begin{aligned}\int_C \mathbf{F}\cdot\mathrm{d}\mathbf{r} &\approx F_u(u,v)h_u(u,v)\,\delta u + F_v(u+\delta u,v)h_v(u+\delta u,v)\,\delta u\\ &\quad - F_u(u,v+\delta v)h_u(u,v+\delta v)\,\delta u - F_v(u,v)h_v(u,v)\,\delta v\\ &\approx \left[\frac{\partial}{\partial u}h_v F_v - \frac{\partial}{\partial v}(h_u F_u)\right]\delta u\delta v.\end{aligned}$$

Divide by the area and take the limit as area $\to 0$, we obtain

$$\lim_{A\to 0}\frac{1}{A}\int_C \mathbf{F}\cdot\mathrm{d}\mathbf{r} = \frac{1}{h_u h_v}\left[\frac{\partial}{\partial u}h_v F_v - \frac{\partial}{\partial v}(h_u F_u)\right].$$

So, by the integral definition of divergence,

$$\mathbf{e}_w\cdot\nabla\times\mathbf{F} = \frac{1}{h_u h_v}\left[\frac{\partial}{\partial u}(h_v F_v) - \frac{\partial}{\partial v}(h_u F_u)\right],$$

and similarly for other components.

We can find the divergence similarly. □

Example. Let $\mathbf{A} = \frac{1}{r}\tan\frac{\theta}{2}\mathbf{e}_\varphi$ in spherical polars. Then

$$\nabla\times\mathbf{A} = \frac{1}{r^2\sin\theta}\begin{vmatrix}\mathbf{e}_r & r\mathbf{e}_\theta & r\sin\theta\mathbf{e}_\varphi\\ \frac{\partial}{\partial r} & \frac{\partial}{\partial\theta} & \frac{\partial}{\partial\varphi}\\ 0 & 0 & r\sin\theta\cdot\frac{1}{r}\tan\frac{\theta}{2}\end{vmatrix} = \frac{\mathbf{e}_r}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left[\sin\theta\tan\frac{\theta}{2}\right] = \frac{1}{r^2}\mathbf{e}_r.$$

10 Gauss' Law and Poisson's equation

10.1 Laws of gravitation

Consider a distribution of mass producing a gravitational force $\mathbf{F}$ on a point mass m at $\mathbf{r}$. The total force is a sum of contributions from each part of the mass distribution, and is proportional to m. Write

$$\mathbf{F} = m\mathbf{g}(\mathbf{r}),$$

Definition (Gravitational field)**.** $\mathbf{g}(\mathbf{r})$ is the *gravitational field*, *acceleration due to gravity*, or *force per unit mass*.

The gravitational field is conservative, ie

$$\oint_C \mathbf{g} \cdot \mathrm{d}\mathbf{r} = 0.$$

This means that if you walk around the place and return to the same position, the total work done is 0 and you did not gain energy, i.e. gravitational potential energy is conserved.

Gauss' law tells us what this gravitational field looks like:

Law (Gauss' law for gravitation)**.** Given any volume V bounded by closed surface S,

$$\int_S \mathbf{g} \cdot \mathrm{d}\mathbf{S} = -4\pi GM,$$

where G is Newton's gravitational constant, and M is the total mass contained in V.

These equations determine $\mathbf{g}(\mathbf{r})$ from a mass distribution.

Example. We can obtain Newton's law of gravitation from Gauss' law together with an assumption about symmetry.

Consider a total mass M distributed with a spherical symmetry about the origin $\mathbf{O}$, with all the mass contained within some radius $r = a$. By spherical symmetry, we have $\mathbf{g}(\mathbf{r}) = g(\mathbf{r})\hat{\mathbf{r}}$.

Consider Gauss' law with S being a sphere of radius $r = R > a$. Then $\hat{\mathbf{n}} = \hat{\mathbf{r}}$. So

$$\int_S \mathbf{g} \cdot \mathrm{d}\mathbf{S} = \int_S g(R)\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}\, \mathrm{d}S = \int g(R)\mathrm{d}S = 4\pi R^2 g(R).$$

By Gauss' law, we obtain

$$4\pi R^2 g(R) = -4\pi GM.$$

So

$$g(R) = -\frac{GM}{R^2}$$

for $R > a$.

Therefore the gravitational force on a mass m at $\mathbf{r}$ is

$$\mathbf{F}(\mathbf{r}) = -\frac{GMm}{r^2}\hat{\mathbf{r}}.$$

If we take the limit as $a \to 0$, we get a point mass M at the origin. Then we recover Newton's law of gravitation for point masses.

The condition $\int_C \mathbf{g} \cdot d\mathbf{r} = 0$ for any closed C can be re-written by Stoke's theorem as

$$\int_S \nabla \times \mathbf{g} \cdot d\mathbf{S} = 0,$$

where S is bounded by the closed curve C. This is true for arbitrary S. So

$$\nabla \times \mathbf{g} = 0.$$

In our example above, $\nabla \times \mathbf{g} = 0$ due to spherical symmetry. But here we showed that it is true for all cases.

Note that we exploited symmetry to solve Gauss' law. However, if the mass distribution is not sufficiently symmetrical, Gauss' law in integral form can be difficult to use. But we can rewrite it in differential form. Suppose

$$M = \int_V \rho(\mathbf{r})\, dV,$$

where ρ is the mass density. Then by Gauss' theorem

$$\int_S \mathbf{g} \cdot d\mathbf{S} = -4\pi GM \Rightarrow \int_V \nabla \cdot \mathbf{g}\, dV = \int_V -4\pi G\rho\, dV.$$

Since this is true for all V, we must have

Law (Gauss' Law for gravitation in differential form)**.**

$$\nabla \cdot \mathbf{g} = -4\pi G\rho.$$

Since $\nabla \times \mathbf{g} = 0$, we can introduce a gravitational potential $\varphi(\mathbf{r})$ with $\mathbf{g} = -\nabla\varphi$. Then Gauss' Law becomes

$$\nabla^2\varphi = 4\pi G\rho.$$

In the example with spherical symmetry, we can solve that

$$\varphi(r) = -\frac{GM}{r}$$

for $r > a$.

10.2 Laws of electrostatics

Consider a distribution of electric charge at rest. They produce a force on a charge q, at rest at $\mathbf{r}$, which is proportional to q.

Definition (Electric field)**.** The force produced by electric charges on another charge q is $\mathbf{F} = q\mathbf{E}(\mathbf{r})$, where $\mathbf{E}(\mathbf{r})$ is the *electric field*, or force per unit charge.

Again, this is conservative. So

$$\oint_C \mathbf{E} \cdot d\mathbf{r} = 0$$

for any closed curve C. It also obeys

Law (Gauss' law for electrostatic forces)**.**

$$\int_S \mathbf{E} \cdot \mathrm{d}\mathbf{S} = \frac{Q}{\varepsilon_0},$$

where ε_0 is the *permittivity of free space*, or *electric constant*.

Then we can write it in differential form, as in the gravitational case.

Law (Gauss' law for electrostatic forces in differential form)**.**

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}.$$

Assuming constant (or no) magnetic field, we have

$$\nabla \times \mathbf{E} = 0.$$

So we can write $\mathbf{E} = -\nabla\varphi$.

Definition (Electrostatic potential)**.** If we write $\mathbf{E} = -\nabla\varphi$, then φ is the *electrostatic potential*, and

$$\nabla^2\varphi = \frac{\rho}{\varepsilon_0}.$$

Example. Take a spherically symmetric charge distribution about O with total charge Q. Suppose all charge is contained within a radius $r = a$. Then similar to the gravitational case, we have

$$\mathbf{E}(\mathbf{r}) = \frac{Q\hat{\mathbf{r}}}{4\pi\varepsilon_0 r^2},$$

and

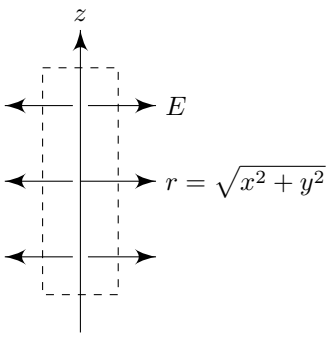
$$\varphi(\mathbf{r}) = \frac{-Q}{4\pi\varepsilon_0 r}.$$

As $a \to 0$, we get *point charges*. From $\mathbf{E}$, we can recover Coulomb's law for the force on another charge q at $\mathbf{r}$:

$$\mathbf{F} = q\mathbf{E} = \frac{qQ\hat{\mathbf{r}}}{4\pi\varepsilon_0 r^2}.$$

Example (Line charge)**.** Consider an infinite line with uniform charge density *per unit length* σ.

We use cylindrical polar coordinates:



By symmetry, the field is radial, i.e.

$$\mathbf{E}(r) = E(r)\hat{\mathbf{r}}.$$

Pick S to be a cylinder of length L and radius r. We know that the end caps do not contribute to the flux since the field lines are perpendicular to the normal. Also, the curved surface has area $2\pi rL$. Then by Gauss' law in integral form,

$$\int_S \mathbf{E} \cdot \mathrm{d}\mathbf{S} = E(r)2\pi rL = \frac{\sigma L}{\varepsilon_0}.$$

So

$$\mathbf{E}(r) = \frac{\sigma}{2\pi\varepsilon_0 r}\hat{\mathbf{r}}.$$

Note that the field varies as $1/r$, not $1/r^2$. Intuitively, this is because we have one more dimension of "stuff" compared to the point charge, so the field does not drop as fast.

10.3 Poisson's Equation and Laplace's equation

Definition (Poisson's equation)**.** The *Poisson's equation* is

$$\nabla^2\varphi = -\rho,$$

where ρ is given and $\varphi(\mathbf{r})$ is to be solved.

This is the form of the equations for gravity and electrostatics, with $-4\pi G\rho$ and ρ/ε_0 in place of ρ respectively.

When $\rho = 0$, we get

Definition (Laplace's equation)**.** Laplace's equation is

$$\nabla^2\varphi = 0.$$

One example is irrotational and incompressible fluid flow: if the velocity is $\mathbf{u}(\mathbf{r})$, then irrotationality gives $\mathbf{u} = \nabla\varphi$ for some *velocity potential* φ. Since it is incompressible, $\nabla \cdot \mathbf{u} = 0$ (cf. previous chapters). So $\nabla^2\varphi = 0$.

The expressions for ∇^2 can be found in non-Cartesian coordinates, but are a bit complicated.

We're concerned here mainly with cases exhibiting spherical or cylindrical symmetry (use r for radial coordinate here). i.e. when $\varphi(\mathbf{r})$ has spherical or cylindrical symmetry. Write $\varphi = \varphi(r)$. Then

$$\nabla\varphi = \varphi'(r)\hat{\mathbf{r}}.$$

Then Laplace's equation $\nabla^2\varphi = 0$ becomes an ordinary differential equation.

– For spherical symmetry, using the chain rule, we have

$$\nabla^2\varphi = \varphi'' + \frac{2}{r}\varphi' = \frac{1}{r^2}(r^2\varphi')' = 0.$$

Then the general solution is

$$\varphi = \frac{A}{r} + B.$$

– For cylindrical symmetry, with $r^2 = x_1^2 + x_2^2$, we have

$$\nabla^2 \varphi = \varphi'' + \frac{1}{r}\varphi' = \frac{1}{r}(r\varphi')' = 0.$$

Then

$$\varphi = A \ln r + B.$$

Then solutions to Poisson's equations can be obtained in a similar way, i.e. by integrating the differential equations directly, or by adding particular integrals to the solutions above.

For example, for a spherically symmetric solution of $\nabla^2 \varphi = -\rho_0$, with ρ_0 constant, recall that $\nabla^2 r^\alpha = \alpha(\alpha+1) r^{\alpha-2}$. Taking $\alpha = 2$, we find the particular integral

$$\varphi = -\frac{\rho_0}{6} r^2,$$

So the general solution with spherical symmetry and constant ρ_0 is

$$\varphi(r) = \frac{A}{r} + B - \frac{1}{6}\rho_0 r^2.$$

To determine A, B, we must specify boundary conditions. If φ is defined on all of $\mathbb{R}^3$, we often require $\varphi \to 0$ as $|\mathbf{r}| \to \infty$. If φ is defined on a bounded volume V, then there are two kinds of common boundary conditions on ∂V:

– Specify φ on ∂V — a *Dirichlet* condition

– Specify $\mathbf{n} \cdot \nabla \varphi$ (sometimes written as $\frac{\partial \varphi}{\partial \mathbf{n}}$): a *Neumann* condition. ($\mathbf{n}$ is the outward normal on ∂V).

The type of boundary conditions we get depends on the physical content of the problem. For example, specifying $\frac{\partial \varphi}{\partial \mathbf{n}}$ corresponds to specifying the normal component of $\mathbf{g}$ or $\mathbf{E}$.

We can also specify different boundary conditions on different boundary components.

Example. We might have a spherically symmetric distribution with constant ρ_0, defined in $a \le r \le b$, with $\varphi(a) = 0$ and $\frac{\partial \varphi}{\partial \mathbf{n}}(b) = 0$.

Then the general solution is

$$\varphi(r) = \frac{A}{r} + B - \frac{1}{6}\rho_0 r^2.$$

We apply the first boundary condition to obtain

$$\frac{A}{a} + B - \frac{1}{6}\rho_0 a^2 = 0.$$

The second boundary condition gives

$$\mathbf{n} \cdot \nabla \varphi = -\frac{A}{b^2} - \frac{1}{3}\rho_0 b = 0.$$

These conditions give

$$A = -\frac{1}{3}\rho_0 b^3, \quad B = \frac{1}{5}\rho_0 a^2 + \frac{1}{3}\rho_0 \frac{b^3}{a}.$$

Example. We might also be interested with spherically symmetric solution with

$$\nabla^2 \varphi = \begin{cases} -\rho_0 & r \le a \\ 0 & r > a \end{cases}$$

with φ non-singular at $r = 0$ and $\varphi(r) \to 0$ as $r \to \infty$, and φ, φ' continuous at $r = a$. This models the gravitational potential on a uniform planet.

Then the general solution from above is

$$\varphi = \begin{cases} \frac{A}{r} + B - \frac{1}{6}\rho_0 r^2 & r \le a \\ \frac{C}{r} + D & r > a. \end{cases}$$

Since φ is non-singular at $r = 0$, we have $A = 0$. Since $\varphi \to 0$ as $r \to \infty$, $D = 0$. So

$$\varphi = \begin{cases} B - \frac{1}{6}\rho_0 r^2 & r \le a \\ \frac{C}{r} & r > a. \end{cases}$$

This is the gravitational potential inside and outside a planet of constant density ρ_0 and radius a. We want φ and φ' to be continuous at $r = a$. So we have

$$B + \frac{1}{6} 4\pi\rho_0 G a^2 = \frac{C}{a}$$
$$\frac{4}{3}\pi G \rho_0 a = -\frac{C}{a^2}.$$

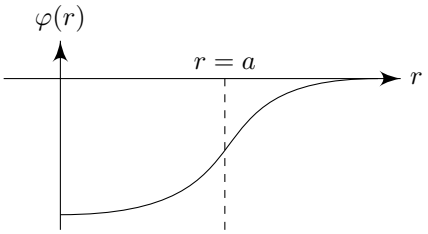
The second equation gives $C = -GM$. Substituting that into the first equation to find B, we get

$$\varphi(r) = \begin{cases} \frac{GM}{2a}\left[\left(\frac{r}{a}\right)^2 - 3\right)\right] & r \le a \\ -\frac{GM}{r} & r > a \end{cases}$$

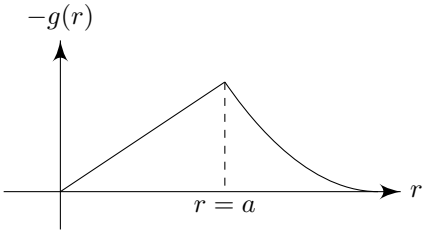
Since $g = -\varphi'$, we have

$$g(r) = \begin{cases} -\frac{GMr}{a^3} & r \le a \\ -\frac{GM}{r} & r > a \end{cases}$$

We can plot the potential energy:



We can also plot $-g(r)$, the inward acceleration:



Alternatively, we can apply Gauss' Law for a flux of $\mathbf{g} = g(r)\mathbf{e}_r$ out of S, a sphere of radius R. For $R \leq a$,

$$\int_S \mathbf{g} \cdot \mathrm{d}\mathbf{S} = 4\pi R^2 g(R) = -4\pi GM \left(\frac{R}{a}\right)^3$$

So

$$g(R) = -\frac{GMR}{a^3}.$$

For $R \geq a$, we can simply apply Newton's law of gravitation.

In general, even if the problem has nothing to do with gravitation or electrostatics, if we want to solve $\nabla^2 \varphi = -\rho$ with ρ and φ sufficiently symmetric, we can consider the flux of $\nabla\varphi$ out of a surface $S = \partial V$:

$$\int_S \nabla\varphi \cdot \mathrm{d}\mathbf{S} = -\int_V \rho \, \mathrm{d}V,$$

by divergence theorem. This is called the Gauss Flux method.

11 Laplace's and Poisson's equations

11.1 Uniqueness theorems

Theorem. Consider $\nabla^2\varphi = -\rho$ for some $\rho(\mathbf{r})$ on a bounded volume V with $S = \partial V$ being a closed surface, with an outward normal $\mathbf{n}$.

Suppose φ satisfies either

(i) Dirichlet condition, $\varphi(\mathbf{r}) = f(\mathbf{r})$ on S

(ii) Neumann condition $\frac{\partial\varphi(\mathbf{r})}{\partial\mathbf{n}} = n\cdot\nabla\varphi = g(\mathbf{r})$ on S.

where f, g are given. Then

(i) $\varphi(\mathbf{r})$ is unique

(ii) $\varphi(\mathbf{r})$ is unique up to a constant.

This theorem is practically important - if you find a solution by any magical means, you know it is the *only* solution (up to a constant).

Since the proof of the cases of the two different boundary conditions are very similar, they will be proved together. When the proof is broken down into (i) and (ii), it refers to the specific cases of each boundary condition.

Proof. Let $\varphi_1(\mathbf{r})$ and $\varphi_2(\mathbf{r})$ satisfy Poisson's equation, each obeying the boundary conditions (N) or (D). Then $\Psi(\mathbf{r}) = \varphi_2(\mathbf{r}) - \varphi_1(\mathbf{r})$ satisfies $\nabla^2\Psi = 0$ on V by linearity, and

(i) $\Psi = 0$ on S; or

(ii) $\frac{\partial\Psi}{\partial\mathbf{n}} = 0$ on S.

Combining these two together, we know that $\Psi\frac{\partial\Psi}{\partial\mathbf{n}} = 0$ on the surface. So using the divergence theorem,

$$\int_V \nabla\cdot(\Psi\nabla\Psi)\,\mathrm{d}V = \int_S (\Psi\nabla\Psi)\cdot\mathrm{d}\mathbf{S} = 0.$$

But

$$\nabla\cdot(\Psi\nabla\Psi) = (\nabla\Psi)\cdot(\nabla\Psi) + \Psi\underbrace{\nabla^2\Psi}_{=0} = |(\nabla\Psi)|^2.$$

So

$$\int_V |\nabla\Psi|^2\,\mathrm{d}V = 0.$$

Since $|\nabla\Psi|^2 \geq 0$, the integral can only vanish if $|\nabla\Psi| = 0$. So $\nabla\Psi = 0$. So $\Psi = c$, a constant on V. So

(i) $\Psi = 0$ on $S \Rightarrow c = 0$. So $\varphi_1 = \varphi_2$ on V.

(ii) $\varphi_2(\mathbf{r}) = \varphi_1(\mathbf{r}) + C$, as claimed. □

We've proven uniqueness. How about existence? It turns out it isn't difficult to craft a boundary condition in which there are no solutions.

For example, if we have $\nabla^2\varphi = -\rho$ on V with the condition $\frac{\partial\varphi}{\partial\mathbf{n}} = g$, then by the divergence theorem,

$$\int_V \nabla^2\varphi \,\mathrm{d}V = \int_{\partial S} \frac{\partial\varphi}{\partial\mathbf{n}} \,\mathrm{d}S.$$

Using Poisson's equation and the boundary conditions, we have

$$\int_V \rho \,\mathrm{d}V + \int_{\partial V} g \,\mathrm{d}S = 0$$

So if ρ and g don't satisfy this equation, then we can't have any solutions.

The theorem can be similarly proved and stated for regions in $\mathbb{R}^2, \mathbb{R}^3, \cdots$, by using the definitions of grad, div and the divergence theorem. The result also extends to unbounded domains. To prove it, we can take a sphere of radius R and impose the boundary conditions $|\Psi(\mathbf{r})| = O(1/R)$ or $|\frac{\partial\Psi}{\partial\mathbf{n}}(\mathbf{r})| = O(1/R^2)$ as $R \to \infty$. Then we just take the relevant limits to complete the proof.

Similar results also apply to related equations and different kinds of boundary conditions, eg D or N on different parts of the boundary. But we have to analyse these case by case and see if the proof still applies.

The proof uses a special case of the result

Proposition (Green's first identity)**.**

$$\int_S (u\nabla v)\cdot \mathrm{d}S = \int_V (\nabla u)\cdot(\nabla v)\,\mathrm{d}V + \int_V u\nabla^2 v \,\mathrm{d}V,$$

By swapping u and v around and subtracting the equations, we have

Proposition (Green's second identity)**.**

$$\int_S (u\nabla v - v\nabla u)\cdot \mathrm{d}\mathbf{S} = \int_V (u\nabla^2 v - v\nabla^2 u)\,\mathrm{d}V.$$

These are sometimes useful, but can be easily deduced from the divergence theorem when needed.

11.2 Laplace's equation and harmonic functions

Definition (Harmonic function)**.** A *harmonic function* is a solution to Laplace's equation $\nabla^2\varphi = 0$.

These have some very special properties.

11.2.1 The mean value property

Proposition (Mean value property)**.** Suppose $\varphi(\mathbf{r})$ is harmonic on region V containing a solid sphere defined by $|\mathbf{r}-\mathbf{a}| \le R$, with boundary $S_R = |\mathbf{r}-\mathbf{a}| = R$, for some R. Define

$$\bar{\varphi}(R) = \frac{1}{4\pi R^2}\int_{S_R} \varphi(\mathbf{r})\,\mathrm{d}S.$$

Then $\varphi(\mathbf{a}) = \bar{\varphi}(R)$.

In words, this says that the value at the center of a sphere is the average of the values on the surface on the sphere.

Proof. Note that $\bar{\varphi}(R) \to \varphi(\mathbf{a})$ as $R \to 0$. We take spherical coordinates (u, θ, χ) centered on $\mathbf{r} = \mathbf{a}$. The scalar element (when $u = R$) on S_R is

$$\mathrm{d}S = R^2 \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\chi.$$

So $\frac{\mathrm{d}S}{R^2}$ is independent of R. Write

$$\bar{\varphi}(R) = \frac{1}{4\pi} \int \varphi \, \frac{\mathrm{d}S}{R^2}.$$

Differentiate this with respect to R, noting that $\mathrm{d}S/R^2$ is independent of R. Then we obtain

$$\frac{\mathrm{d}}{\mathrm{d}R}\bar{\varphi}(R) = \frac{1}{4\pi R^2} \int \left.\frac{\partial \varphi}{\partial u}\right|_{u=R} \mathrm{d}S$$

But

$$\frac{\partial \varphi}{\partial u} = \mathbf{e}_u \cdot \nabla \varphi = \mathbf{n} \cdot \nabla \varphi = \frac{\partial \varphi}{\partial \mathbf{n}}$$

on S_R. So

$$\frac{\mathrm{d}}{\mathrm{d}R}\bar{\varphi}(R) = \frac{1}{4\pi R^2} \int_{S_R} \nabla \varphi \cdot \mathrm{d}\mathbf{S} = \frac{1}{4\pi R^2} \int_{V_R} \nabla^2 \varphi \, \mathrm{d}V = 0$$

by divergence theorem. So $\bar{\varphi}(R)$ does not depend on R, and the result follows. □

11.2.2 The maximum (or minimum) principle

In this section, we will talk about maxima of functions. It should be clear that the results also hold for minima.

Definition (Local maximum)**.** We say that $\varphi(\mathbf{r})$ has a *local maximum* at $\mathbf{a}$ if for some $\varepsilon > 0$, $\varphi(\mathbf{r}) < \varphi(\mathbf{a})$ when $0 < |\mathbf{r} - \mathbf{a}| < \varepsilon$.

Proposition (Maximum principle)**.** If a function φ is harmonic on a region V, then φ cannot have a maximum at an interior point of $\mathbf{a}$ of V.

Proof. Suppose that φ had a local maximum at $\mathbf{a}$ in the interior. Then there is an ε such that for any $\mathbf{r}$ such that $0 < |\mathbf{r} - \mathbf{a}| < \varepsilon$, we have $\varphi(\mathbf{r}) < \varphi(\mathbf{a})$.

Note that if there is an ε that works, then any smaller ε will work. Pick an ε sufficiently small such that the region $|\mathbf{r} - \mathbf{a}| < \varepsilon$ lies within V (possible since $\mathbf{a}$ lies in the interior of V).

Then for any $\mathbf{r}$ such that $|\mathbf{r} - \mathbf{a}| = \varepsilon$, we have $\varphi(\mathbf{r}) < \varphi(\mathbf{a})$.

$$\bar{\varphi}(\varepsilon) = \frac{1}{4\pi R^2} \int_{S_R} \varphi(\mathbf{r}) \, \mathrm{d}S < \varphi(\mathbf{a}),$$

which contradicts the mean value property. □

We can understand this by performing a local analysis of stationary points by differentiation. Suppose at $\mathbf{r} = \mathbf{a}$, we have $\nabla \varphi = 0$. Let the eigenvalues of the Hessian matrix $H_{ij} = \frac{\partial^2}{\partial x_i \partial x_j}$ be λ_i. But since φ is harmonic, we have $\nabla^2 \varphi = 0$,

i.e. $\frac{\partial^2 \varphi}{\partial x_i \partial x_i} = H_{ii} = 0$. But H_{ii} is the trace of the Hessian matrix, which is the sum of eigenvalues. So $\sum \lambda_i = 0$.

Recall that a maximum or minimum occurs when all eigenvalues have the same sign. This clearly cannot happen if the sum is 0. Therefore we can only have saddle points.

(note we ignored the case where all $\lambda_i = 0$, where this analysis is inconclusive)

11.3 Integral solutions of Poisson's equations

11.3.1 Statement and informal derivation

We want to find a solution to Poisson's equations. We start with a discrete case, and try to generalize it to a continuous case.

If there is a single point source of strength λ at $\mathbf{a}$, the potential φ is

$$\varphi = \frac{\lambda}{4\pi} \frac{1}{|\mathbf{r} - \mathbf{a}|}.$$

(we have $\lambda = -4\pi GM$ for gravitation and Q/ε_0 for electrostatics)

If we have many sources λ_α at positions $\mathbf{r}_\alpha$, the potential is a sum of terms

$$\varphi(\mathbf{r}) = \sum_\alpha \frac{1}{4\pi} \frac{\lambda_\alpha}{|\mathbf{r} - \mathbf{r}_\alpha|}.$$

If we have infinitely many of them, having a distribution of $\rho(\mathbf{r})$ with $\rho(\mathbf{r}')\,\mathrm{d}V'$ being the contribution from a small volume at position $\mathbf{r}'$. It would be reasonable to guess that the solution is what we obtain by replacing the sum with an integral:

Proposition. The solution to Poisson's equation $\nabla^2 \varphi = -\rho$, with boundary conditions $|\varphi(\mathbf{r})| = O(1/|\mathbf{r}|)$ and $|\nabla \varphi(\mathbf{r})| = O(1/|\mathbf{r}|^2)$, is

$$\varphi(\mathbf{r}) = \frac{1}{4\pi} \int_{V'} \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\,\mathrm{d}V'$$

For $\rho(\mathbf{r}')$ non-zero everywhere, but suitably well-behaved as $|\mathbf{r}'| \to \infty$, we can also take $V' = \mathbb{R}^3$.

Example. Suppose

$$\nabla^2 \varphi = \begin{cases} -\rho_0 & |\mathbf{r}| \le a \\ 0 & |\mathbf{r}| > a. \end{cases}$$

Fix $\mathbf{r}$ and introduce polar coordinates r', θ, χ for $\mathbf{r}'$. We take the $\theta = 0$ direction to be the direction along the line from $\mathbf{r}'$ to $\mathbf{r}$.

Then

$$\varphi(\mathbf{r}) = \frac{1}{4\pi} \int_{V'} \frac{\rho_0}{|\mathbf{r} - \mathbf{r}'|}\,\mathrm{d}V'.$$

We have

$$\mathrm{d}V' = r'^2 \sin\theta\,\mathrm{d}r'\,\mathrm{d}\theta\,\mathrm{d}\chi.$$

We also have

$$|\mathbf{r} - \mathbf{r}'| = \sqrt{r^2 + r'^2 - 2rr'\cos\theta}$$

by the cosine rule ($c^2 = a^2 + b^2 - 2ab\cos C$). So

$$\begin{aligned}\varphi(\mathbf{r}) &= \frac{1}{4\pi}\int_0^a \mathrm{d}r' \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\chi \frac{\rho_0 r'^2 \sin\theta}{\sqrt{r^2 + r'^2 - 2rr'\cos\theta}} \\ &= \frac{\rho_0}{2}\int_0^a \mathrm{d}r' \frac{r'^2}{rr'}\left[\sqrt{r^2 + r'^2 - rr'\cos\theta}\right]_{\theta=0}^{\theta=\pi} \\ &= \frac{\rho_0}{2}\int_0^a \mathrm{d}r' \frac{r'}{r}(|\mathbf{r}+\mathbf{r}'| + |\mathbf{r}-\mathbf{r}'|) \\ &= \frac{\rho_0}{2}\int_0^a \left[\mathrm{d}r' \frac{r'}{r}\left(\begin{cases} 2r' & r > r' \\ 2r & r < r' \end{cases}\right)\right]\end{aligned}$$

If $r > a$, then $r > r'$ always. So

$$\varphi(\mathbf{r}) = \rho_0 \int_0^a \frac{r'^2}{r}\,\mathrm{d}r' = \frac{r_0 a^3}{3r}.$$

If $r < a$, then the integral splits into two parts:

$$\varphi(\mathbf{r}) = \rho_0\left(\int_0^r \mathrm{d}r' \frac{r'^2}{r} + \int_r^a \mathrm{d}r' r'\right) = \rho_0\left[-\frac{1}{6}r^2 + \frac{a^2}{2}\right].$$

11.3.2 Point sources and δ-functions*

Recall that

$$\Psi = \frac{\lambda}{4\pi|\mathbf{r}-\mathbf{a}|}$$

is our potential for a point source. When $\mathbf{r} \neq \mathbf{a}$, we have

$$\nabla\Psi = -\frac{\lambda}{4\pi}\frac{\mathbf{r}-\mathbf{a}}{|\mathbf{r}-\mathbf{a}|^3}, \quad \nabla^2\Psi = 0.$$

What about when $\mathbf{r} = \mathbf{a}$? Ψ is singular at this point, but can we say anything about $\nabla^2\Psi$?

For any sphere with center $\mathbf{a}$, we have

$$\int_S \nabla\Psi \cdot \mathrm{d}\mathbf{S} = -\lambda.$$

By the divergence theorem, we have

$$\int \nabla^2\Psi \,\mathrm{d}V = -\lambda.$$

for V being a solid sphere with $\partial V = S$. Since $\nabla^2\Psi$ is zero at any point $\mathbf{r} \neq \mathbf{a}$, we must have

$$\nabla^2\Psi = -\lambda\delta(\mathbf{r}-\mathbf{a}),$$

where δ is the 3d delta function, which satisfies

$$\int_V f(\mathbf{r})\delta(\mathbf{r}-\mathbf{a})\,\mathrm{d}V = f(\mathbf{a})$$

for any volume containing $\mathbf{a}$.

In short, we have

$$\nabla^2 \left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) = -4\pi\delta(\mathbf{r} - \mathbf{r}').$$

Using these, we can verify that the integral solution of Poisson's equation we obtained previously is correct:

$$\begin{aligned}
\nabla^2\Psi(\mathbf{r}) &= \nabla^2 \left(\frac{1}{4\pi} \int_{V'} \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, \mathrm{d}V' \right) \\
&= \frac{1}{4\pi} \int_{V'} \rho(\mathbf{r}') \nabla^2 \left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) \, \mathrm{d}V' \\
&= -\int_{V'} \rho(\mathbf{r}')\delta(\mathbf{r} - \mathbf{r}') \, \mathrm{d}V' \\
&= -\rho(\mathbf{r}),
\end{aligned}$$

as required.

12 Maxwell's equations

12.1 Laws of electromagnetism

Maxwell's equations are a set of four equations that describe the behaviours of electromagnetism. Together with the Lorentz force law, these describe *all* we know about (classical) electromagnetism. All other results we know are simply mathematical consequences of these equations. It is thus important to understand the mathematical properties of these equations.

To begin with, there are two fields that govern electromagnetism, known as the *electric* and *magnetic* field. These are denoted by $\mathbf{E}(r,t)$ and $\mathbf{B}(r,t)$ respectively.

To understand electromagnetism, we need to understand how these fields are formed, and how these fields affect charged particles. The second is rather straightforward, and is given by the Lorentz force law.

Law (Lorentz force law)**.** A point charge q experiences a force of

$$\mathbf{F} = q(\mathbf{E} + \dot{\mathbf{r}} \times \mathbf{B}).$$

The dynamics of the field itself is governed by *Maxwell's equations*. To state the equations, we need to introduce two more concepts.

Definition (Charge and current density)**.** $\rho(\mathbf{r}, t)$ is the *charge density*, defined as the charge per unit volume.

$\mathbf{j}(\mathbf{r}, t)$ is the *current density*, defined as the electric current per unit area of cross section.

Then Maxwell's equations say

Law (Maxwell's equations)**.**

$$\begin{aligned} \nabla \cdot \mathbf{E} &= \frac{\rho}{\varepsilon_0} \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} &= 0 \\ \nabla \times \mathbf{B} - \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} &= \mu_0 \mathbf{j}, \end{aligned}$$

where ε_0 is the electric constant (permittivity of free space) and μ_0 is the magnetic constant (permeability of free space), which are constants determined experimentally.

We can quickly derive some properties we know from these four equations. The conservation of electric charge comes from taking the divergence of the last equation.

$$\underbrace{\nabla \cdot (\nabla \times \mathbf{B})}_{=0} - \mu_0 \varepsilon_0 \frac{\partial}{\partial t} \underbrace{(\nabla \cdot \mathbf{E})}_{=\rho/\varepsilon_0} = \mu_0 \nabla \cdot \mathbf{j}.$$

So

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0.$$

We can also take the volume integral of the first equation to obtain

$$\int_V \nabla \cdot \mathbf{E} \, \mathrm{d}V = \frac{1}{\varepsilon_0} \int_V \rho \, \mathrm{d}V = \frac{Q}{\varepsilon_0}.$$

By the divergence theorem, we have

$$\int_S \mathbf{E} \cdot \mathrm{d}\mathbf{S} = \frac{Q}{\varepsilon_0},$$

which is Gauss' law for electric fields

We can integrate the second equation to obtain

$$\int_S \mathbf{B} \cdot \mathrm{d}\mathbf{S} = 0.$$

This roughly states that there are no "magnetic charges".

The remaining Maxwell's equations also have integral forms. For example,

$$\int_{C=\partial S} \mathbf{E} \cdot \mathrm{d}\mathbf{r} = \int_S \nabla \times \mathbf{E} \, \mathrm{d}S = -\frac{\mathrm{d}}{\mathrm{d}t} \int_S \mathbf{B} \cdot \mathrm{d}\mathbf{S},$$

where the first equality is from from Stoke's theorem. This says that a changing magnetic field produces a current.

12.2 Static charges and steady currents

If $\rho, \mathbf{j}, \mathbf{E}, \mathbf{B}$ are all independent of time, $\mathbf{E}$ and $\mathbf{B}$ are no longer linked.

We can solve the equations for electric fields:

$$\nabla \cdot \mathbf{E} = \rho/\varepsilon_0$$
$$\nabla \times \mathbf{E} = \mathbf{0}$$

Second equation gives $\mathbf{E} = -\nabla\varphi$. Substituting into first gives $\nabla^2\varphi = -\rho/\varepsilon_0$.

The equations for the magnetic field are

$$\nabla \cdot \mathbf{B} = 0$$
$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j}$$

First equation gives $B = \nabla \times \mathbf{A}$ for some *vector potential* $\mathbf{A}$. But the vector potential is not well-defined. Making the transformation $\mathbf{A} \mapsto \mathbf{A} + \nabla\chi(\mathbf{x})$ produces the same $\mathbf{B}$, since $\nabla \times (\nabla\chi) = 0$. So choose χ such that $\nabla \cdot \mathbf{A} = 0$. Then

$$\nabla^2 \mathbf{A} = \nabla(\underbrace{\nabla \cdot \mathbf{A}}_{=0}) - \nabla \times (\underbrace{\nabla \times \mathbf{A}}_{\mathbf{B}}) = -\mu_0 \mathbf{j}.$$

In summary, we have

Electrostatics	Magnetostatics
$\nabla \cdot \mathbf{E} = \rho/\varepsilon_0$ $\nabla \times \mathbf{E} = \mathbf{0}$ $\nabla^2\varphi = -\rho/\varepsilon_0$ ε_0 sets the scale of electrostatic effects, e.g. the Coulomb force	$\nabla \cdot \mathbf{B} = 0$ $\nabla \times \mathbf{B} = \mu_0 \mathbf{j}$ $\nabla^2 \mathbf{A} = -\mu_0 \mathbf{j}$. μ_0 sets the scale of magnetic effects, e.g. force between two wires with currents.

12.3 Electromagnetic waves

Consider Maxwell's equations in empty space, i.e. $\rho = 0$, $\mathbf{j} = \mathbf{0}$. Then Maxwell's equations give

$$\nabla^2 \mathbf{E} = \nabla(\nabla \cdot \mathbf{E}) - \nabla \times (\nabla \times \mathbf{E}) = \nabla \times \frac{\partial \mathbf{B}}{\partial t} = \frac{\partial}{\partial t}(\nabla \times \mathbf{B}) = \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial^2 t}.$$

Define $c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$. Then the equation gives

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\mathbf{E} = 0.$$

This is the wave equation describing propagation with speed c. Similarly, we can obtain

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\mathbf{B} = 0.$$

So Maxwell's equations predict that there exists electromagnetic waves in free space, which move with speed $c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} \approx 3.00 \times 10^8\,\mathrm{m\,s^{-1}}$, which is the speed of light! Maxwell then concluded that light is electromagnetic waves!

13 Tensors and tensor fields

13.1 Definition

There are two ways we can think of a vector in $\mathbb{R}^3$. We can either interpret it as a "point" in space, or we can view it simply as a list of three numbers. However, the list of three numbers is simply a *representation* of the vector with respect to some particular basis. When we change basis, in order to represent the same point, we will need to use a different list of three numbers. In particular, when we perform a rotation by R_{ip}, the new components of the vector is given by

$$v_i' = R_{ip} v_p.$$

Similarly, we can imagine a matrix as either a linear transformation or an array of 9 numbers. Again, when we change basis, in order to represent the same transformation, we will need a different array of numbers. This time, the transformation is given by

$$A_{ij}' = R_{ip} R_{jq} A_{pq}.$$

We can think about this from another angle. To define an arbitrary quantity A_{ij}, we can always just write down 9 numbers and be done with it. Moreover, we can write down a different set of numbers in a different basis. For example, we can define $A_{ij} = \delta_{ij}$ in our favorite basis, but $A_{ij} = 0$ in all other bases. We can do so because we have the power of the pen.

However, for this A_{ij} to represent something physically meaningful, i.e. an actual linear transformation, we have to make sure that the components of A_{ij} transform sensibly under a basis transformation. By "sensibly", we mean that it has to follow the transformation rule $A_{ij}' = R_{ip} R_{jq} A_{pq}$. For example, the A_{ij} we defined in the previous paragraph does *not* transform sensibly. While it is something we can define and write down, it does not correspond to anything meaningful.

The things that transform sensibly are known as *tensors*. For example, vectors and matrices (that transform according to the usual change-of-basis rules) are tensors, but that A_{ij} is not.

In general, tensors are allowed to have an arbitrary number of indices. In order for a quantity $T_{ij\cdots k}$ to be a tensor, we require it to transform according to

$$T_{ij\cdots k}' = R_{ip} R_{jq} \cdots R_{kr} T_{pq\cdots r},$$

which is an obvious generalization of the rules for vectors and matrices.

Definition (Tensor)**.** A *tensor* of rank n has components $T_{ij\cdots k}$ (with n indices) with respect to each basis $\{\mathbf{e}_i\}$ or coordinate system $\{x_i\}$, and satisfies the following rule of change of basis:

$$T_{ij\cdots k}' = R_{ip} R_{jq} \cdots R_{kr} T_{pq\cdots r}.$$

Example.

- A tensor T of rank 0 doesn't transform under change of basis, and is a scalar.

– A tensor T of rank 1 transforms under $T'_i = R_{ip}T_p$. This is a vector.

– A tensor T of rank 2 transforms under $T'_{ij} = R_{ip}R_{jq}T_{pq}$. This is a matrix.

Example.

(i) If $\mathbf{u}, \mathbf{v}, \cdots \mathbf{w}$ are n vectors, then

$$T_{ij\cdots k} = u_i v_j \cdots w_k$$

defines a tensor of rank n. To check this, we check the tensor transformation rule. We do the case for $n = 2$ for simplicity of expression, and it should be clear that this can be trivially extended to arbitrary n:

$$\begin{aligned} T'_{ij} &= u'_i v'_j = (R_{ip}u_p)(R_{jq}v_q) \\ &= R_{ip}R_{jq}(u_p v_q) \\ &= R_{ip}R_{jq}T_{pq} \end{aligned}$$

Then linear combinations of such expressions are also tensors, e.g. $T_{ij} = u_i v_j + a_i b_j$ for any $\mathbf{u}, \mathbf{v}, \mathbf{a}, \mathbf{b}$.

(ii) δ_{ij} and ε_{ijk} are tensors of rank 2 and 3 respectively — with the special property that their components are unchanged with respect to the basis coordinate:

$$\delta'_{ij} = R_{ip}R_{jq}\delta_{pq} = R_{ip}R_{jp} = \delta_{ij},$$

since $R_{ip}R_{jp} = (RR^T)_{ij} = I_{ij}$. Also

$$\varepsilon'_{ijk} = R_{ip}R_{jq}R_{kr}\varepsilon_{pqr} = (\det R)\varepsilon_{ijk} = \varepsilon_{ijk},$$

using results from Vectors and Matrices.

(iii) (Physical example) In some substances, an applied electric field $\mathbf{E}$ gives rise to a current density $\mathbf{j}$, according to the linear relation $j_i = \varepsilon_{ij}E_j$, where ε_{ij} is the *conductivity tensor*.

Note that this relation entails that the resulting current need not be in the same direction as the electric field. This might happen if the substance has special crystallographic directions that favours electric currents.

However, if the substance is *isotropic*, we have $\varepsilon_{ij} = \sigma\delta_{ij}$ for some σ. In this case, the current *is* parallel to the field.

13.2 Tensor algebra

Definition (Tensor addition)**.** Tensors T and S of the same rank can be *added*; $T + S$ is also a tensor of the same rank, defined as

$$(T + S)_{ij\cdots k} = T_{ij\cdots k} + S_{ij\cdots k}.$$

in any coordinate system.

To check that this is a tensor, we check the transformation rule. Again, we only show for $n = 2$:

$$(T + S)'_{ij} = T'_{ij} + S'_{ij} = R_{ip}R_{jq}T_{pq} + R_{ip}R_{jq}S_{pq} = (R_{ip}R_{jq})(T_{pq} + S_{pq}).$$

Definition (Scalar multiplication)**.** A tensor T of rank n can be multiplied by a scalar α. αT is a tensor of the same rank, defined by

$$(\alpha T)_{ij} = \alpha T_{ij}.$$

It is trivial to check that the resulting object is indeed a tensor.

Definition (Tensor product)**.** Let T be a tensor of rank n and S be a tensor of rank m. The *tensor product* $T \otimes S$ is a tensor of rank $n + m$ defined by

$$(T \otimes S)_{x_1 x_2 \cdots x_n y_1 y_2 \cdots y_m} = T_{x_1 x_2 \cdots x_n} S_{y_1 y_2 \cdots y_n}.$$

It is trivial to show that this is a tensor.

We can similarly define tensor products for any (positive integer) number of tensors, e.g. for n vectors $\mathbf{u}, \mathbf{v} \cdots, \mathbf{w}$, we can define

$$T = \mathbf{u} \otimes \mathbf{v} \otimes \cdots \otimes \mathbf{w}$$

by

$$T_{ij\ldots k} = u_i v_j \cdots w_k,$$

as defined in the example in the beginning of the chapter.

Definition (Tensor contraction)**.** For a tensor T of rank n with components $T_{ijp\cdots q}$, we can *contract on* the indices i, j to obtain a new tensor of rank $n - 2$:

$$S_{p\cdots q} = \delta_{ij} T_{ijp\cdots q} = T_{iip\cdots q}$$

Note that we don't have to always contract on the first two indices. We can contract any pair we like.

To check that contraction produces a tensor, we take the ranks 2 T_{ij} example. Contracting, we get T_{ii} ,a rank-0 scalar. We have $T'_{ii} = R_{ip} R_{iq} T_{pq} = \delta_{pq} T_{pq} = T_{pp} = T_{ii}$, since R is an orthogonal matrix.

If we view T_{ij} as a matrix, then the contraction is simply the trace of the matrix. So our result above says that the trace is invariant under basis transformations — as we already know in IA Vectors and Matrices.

Note that our usual matrix product can be formed by first applying a tensor product to obtain $M_{ij} N_{pq}$, then contract with δ_{jp} to obtain $M_{ij} N_{jq}$.

13.3 Symmetric and antisymmetric tensors

Definition (Symmetric and anti-symmetric tensors)**.** A tensor T of rank n is *symmetric* in the indices i, j if it obeys

$$T_{ijp\cdots q} = T_{jip\cdots q}.$$

It is *anti-symmetric* if

$$T_{ijp\cdots q} = -T_{jip\cdots q}.$$

Again, a tensor can be symmetric or anti-symmetric in any pair of indices, not just the first two.

This is a property that holds in any coordinate systems, if it holds in one, since

$$T'_{k\ell r\dots s} = R_{ki}R_{\ell j}R_{rp}\cdots R_{sq}T_{ijp\cdots q} = \pm R_{ki}R_{\ell j}R_{rp}\cdots R_{sq}T_{jip\cdots q} = \pm T'_{\ell kr\cdots s}$$

as required.

Definition (Totally symmetric and anti-symmetric tensors)**.** A tensor is *totally (anti-)symmetric* if it is (anti-)symmetric in every pair of indices.

Example. $\delta_{ij} = \delta_{ji}$ is totally symmetric, while $\varepsilon_{ijk} = -\varepsilon_{jik}$ is totally antisymmetric.

There are totally symmetric tensors of arbitrary rank n. But in $\mathbb{R}^3$,

- Any totally antisymmetric tensor of rank 3 is $\lambda\varepsilon_{ijk}$ for some scalar λ.
- There are no totally antisymmetric tensors of rank greater than 3, except for the trivial tensor with all components 0.

  Proof: exercise (hint: pigeonhole principle)

13.4 Tensors, multi-linear maps and the quotient rule

Tensors as multi-linear maps

In Vectors and Matrices, we know that matrices are linear maps. We will prove an analogous fact for tensors.

Definition (Multilinear map)**.** A map T that maps n vectors $\mathbf{a}, \mathbf{b}, \cdots, \mathbf{c}$ to $\mathbb{R}$ is multi-linear if it is linear in each of the vectors $\mathbf{a}, \mathbf{b}, \cdots, \mathbf{c}$ individually.

We will show that a tensor T of rank n is a equivalent to a multi-linear map from n vectors $\mathbf{a}, \mathbf{b}, \cdots, \mathbf{c}$ to $\mathbb{R}$ defined by

$$T(\mathbf{a}, \mathbf{b}, \cdots, \mathbf{c}) = T_{ij\dots k}a_i b_j \cdots c_k.$$

To show that tensors are *equivalent* to multi-linear maps, we have to show the following:

(i) Defining a map with a tensor makes sense, i.e. the expression $T_{ij\dots k}a_i b_j \cdots c_k$ is the same regardless of the basis chosen;

(ii) While it is always possible to write a multi-linear map as $T_{ij\dots k}a_i b_j \cdots c_k$, we have to show that $T_{ij\dots k}$ is indeed a tensor, i.e. transform according to the tensor transformation rules.

To show the first property, just note that the $T_{ij\dots k}a_i b_j \cdots c_k$ is a tensor product (followed by contraction), which retains tensor-ness. So it is also a tensor. In particular, it is a rank 0 tensor, i.e. a scalar, which is independent of the basis.

To show the second property, assuming that T is a multi-linear map, it must be independent of the basis, so

$$T_{ij\dots k}a_i b_j \cdots c_k = T'_{ij\dots k}a'_i b'_j \cdots c'_k.$$

Since $v'_p = R_{pi} v_i$ by tensor transformation rules, multiplying both sides by R_{pi} gives $v_i = R_{pi} v'_p$. Substituting in gives

$$T_{ij\cdots k}(R_{pi}a'_p)(R_{qj}b'_q)\cdots(R_{kr}c'_r) = T'_{pq\cdots r}a'_p b'_q \cdots c'_r.$$

Since this is true for all $\mathbf{a}, \mathbf{b}, \cdots \mathbf{c}$, we must have

$$T_{ij\cdots k}R_{pi}R_{qj}\cdots R_{rk} = T'_{pq\cdots r}$$

Hence $T_{ij\cdots k}$ obeys the tensor transformation rule, and is a tensor.

This shows that there is a one-to-one correspondence between tensors of rank n and multi-linear maps.

This gives a way of thinking about tensors independent of any coordinate system or choice of basis, and the tensor transformation rule emerges naturally.

Note that the above is exactly what we did with linear maps and matrices.

The quotient rule

If $T_{\underbrace{i\cdots j}_{n}\underbrace{p\cdots q}_{m}}$ is a tensor of rank $n+m$, and $u_{p\cdots q}$ is a tensor of rank m then

$$v_{i,\cdots j} = T_{i\cdots jp\cdots q}u_{p\cdots q}$$

is a tensor of rank n, since it is a tensor product of T and u, followed by contraction.

The converse is also true:

Proposition (Quotient rule)**.** Suppose that $T_{i\cdots jp\cdots q}$ is an array defined in each coordinate system, and that $v_{i\cdots j} = T_{i\cdots jp\cdots q}u_{p\cdots q}$ is also a tensor for any tensor $u_{p\cdots q}$. Then $T_{i\cdots jp\cdots q}$ is also a tensor.

Note that we have previously seen the special case of $n = m = 1$, which says that linear maps are tensors.

Proof. We can check the tensor transformation rule directly. However, we can reuse the result above to save some writing.

Consider the special form $u_{p\cdots q} = c_p \cdots d_q$ for any vectors $\mathbf{c}, \cdots \mathbf{d}$. By assumption,

$$v_{i\cdots j} = T_{i\cdots jp\cdots q}c_p \cdots d_q$$

is a tensor. Then

$$v_{i\cdots j}a_i \cdots b_j = T_{i\cdots jp\cdots q}a_i \cdots b_j c_p \cdots d_q$$

is a scalar for any vectors $\mathbf{a}, \cdots, \mathbf{b}, \mathbf{c}, \cdots, \mathbf{d}$. Since $T_{i\cdots jp\cdots q}a_i \cdots b_j c_p \cdots d_q$ is a scalar and hence gives the same result in every coordinate system, $T_{i\cdots jp\cdots q}$ is a multi-linear map. So $T_{i\cdots jp\cdots q}$ is a tensor. □

13.5 Tensor calculus

Tensor fields and derivatives

Just as with scalars or vectors, we can define tensor fields:

Definition (Tensor field)**.** A *tensor field* is a tensor at each point in space $T_{ij\cdots k}(\mathbf{x})$, which can also be written as $T_{ij\cdots k}(x_\ell)$.

We assume that the fields are *smooth* so they can be differentiated any number of times

$$\frac{\partial}{\partial x_p}\cdots\frac{\partial}{\partial x_q}T_{ij\cdots k},$$

except for where things obviously fail, e.g. for where T is not defined. We now claim:

Proposition.

$$\underbrace{\frac{\partial}{\partial x_p}\cdots\frac{\partial}{\partial x_q}}_{m}T_{\underbrace{ij\cdots k}_{n}}, \tag{$*$}$$

is a tensor of rank $n+m$.

Proof. To show this, it suffices to show that $\frac{\partial}{\partial x_p}$ satisfies the tensor transformation rules for rank 1 tensors (i.e. it is something like a rank 1 tensor). Then by the exact same argument we used to show that tensor products preserve tensorness, we can show that the $(*)$ is a tensor. (we cannot use the result of tensor products directly, since this is not exactly a product. But the exact same proof works!)

Since $x_i' = R_{iq}x_q$, we have

$$\frac{\partial x_i'}{\partial x_p} = R_{ip}.$$

(noting that $\frac{\partial x_p}{\partial x_q} = \delta_{pq}$). Similarly,

$$\frac{\partial x_q}{\partial x_i'} = R_{iq}.$$

Note that R_{ip}, R_{iq} are constant matrices.

Hence by the chain rule,

$$\frac{\partial}{\partial x_i'} = \left(\frac{\partial x_q}{\partial x_i'}\right)\frac{\partial}{\partial x_q} = R_{iq}\frac{\partial}{\partial x_q}.$$

So $\frac{\partial}{\partial x_p}$ obeys the vector transformation rule. So done. □

Integrals and the tensor divergence theorem

It is also straightforward to do integrals. Since we can sum tensors and take limits, the definition of a tensor-valued integral is straightforward.

For example, $\int_V T_{ij\cdots k}(\mathbf{x})\,\mathrm{d}V$ is a tensor of the same rank as $T_{ij\cdots k}$ (think of the integral as the limit of a sum).

For a physical example, recall our discussion of the flux of quantities for a fluid with velocity $\mathbf{u}(\mathbf{x})$ through a surface element — assume a uniform density ρ. The flux of volume is $\mathbf{u}\cdot\mathbf{n}\delta s = u_j n_j \delta S$. So the flux of mass is $\rho u_j n_j \delta S$. Then the flux of the ith component of momentum is $\rho u_i u_j n_j \delta S = T_{ij} n_j k \delta S$

(mass times velocity), where $T_{ij} = \rho u_i u_j$. Then the flux through the surface S is $\int_S T_{ij} n_j \, \mathrm{d}S$.

It is easy to generalize the divergence theorem from vectors to tensors. We can then use it to discuss conservation laws for tensor quantities.

Let V be a volume bounded by a surface $S = \partial V$ and $T_{ij\cdots k\ell}$ be a smooth tensor field. Then

Theorem (Divergence theorem for tensors)**.**

$$\int_S T_{ij\cdots k\ell} n_\ell \, \mathrm{d}S = \int_V \frac{\partial}{\partial x_\ell}(T_{ij\cdots k\ell}) \, \mathrm{d}V,$$

with $\mathbf{n}$ being an outward pointing normal.

The regular divergence theorem is the case where T has one index and is a vector field.

Proof. Apply the usual divergence theorem to the vector field $\mathbf{v}$ defined by $v_\ell = a_i b_j \cdots c_k T_{ij\cdots k\ell}$, where $\mathbf{a}, \mathbf{b}, \cdots, \mathbf{c}$ are fixed constant vectors.

Then

$$\nabla \cdot \mathbf{v} = \frac{\partial v_\ell}{\partial x_\ell} = a_i b_j \cdots c_k \frac{\partial}{\partial x_\ell} T_{ij\cdots k\ell},$$

and

$$\mathbf{n} \cdot \mathbf{v} = n_\ell v_\ell = a_i b_j \cdots c_k T_{ij\cdots k\ell} n_\ell.$$

Since $\mathbf{a}, \mathbf{b}, \cdots, \mathbf{c}$ are arbitrary, therefore they can be eliminated, and the tensor divergence theorem follows. □

14 Tensors of rank 2

14.1 Decomposition of a second-rank tensor

This decomposition might look arbitrary at first sight, but as time goes on, you will find that it is actually very useful in your future career (at least, the lecturer claims so).

Any second rank tensor can be written as a sum of its symmetric and anti-symmetric parts

$$T_{ij} = S_{ij} + A_{ij},$$

where

$$S_{ij} = \frac{1}{2}(T_{ij} + T_{ji}), \quad A_{ij} = \frac{1}{2}(T_{ij} - T_{ji}).$$

Here T_{ij} has 9 independent components, whereas S_{ij} and A_{ij} have 6 and 3 independent components, since they must be of the form

$$(S_{ij}) = \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix}, \quad (A_{ij}) = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}.$$

The symmetric part can be be further reduced to a *traceless* part plus an *isotropic* (i.e. multiple of δ_{ij}) part:

$$S_{ij} = P_{ij} + \frac{1}{3}\delta_{ij}Q,$$

where $Q = S_{ii}$ is the trace of S_{ij} and $P_{ij} = P_{ji} = S_{ij} - \frac{1}{3}\delta_{ij}Q$ is traceless. Then P_{ij} has 5 independent components while Q has 1.

Since the antisymmetric part has 3 independent components, just like a usual vector, we should be able to write A_i in terms of a single vector. In fact, we can write the antisymmetric part as

$$A_{ij} = \varepsilon_{ijk}B_k$$

for some vector B. To figure out what this B is, we multiply by $\varepsilon_{ij\ell}$ on both sides and use some magic algebra to obtain

$$B_k = \frac{1}{2}\varepsilon_{ijk}A_{ij} = \frac{1}{2}\varepsilon_{ijk}T_{ij},$$

where the last equality is from the fact that only antisymmetric parts contribute to the sum.

Then

$$(A_{ij}) = \begin{pmatrix} 0 & B_3 & -B_2 \\ -B_3 & 0 & B_1 \\ B_2 & -B_1 & 0 \end{pmatrix}$$

To summarize,

$$T_{ij} = P_{ij} + \varepsilon_{ijk}B_k + \frac{1}{3}\delta_{ij}Q,$$

where $B_k = \frac{1}{2}\varepsilon_{pqj}T_{pq}$, $Q = T_{kk}$ and $P_{ij} = P_{ji} = \frac{T_{ij}+T_{ji}}{2} - \frac{1}{3}\delta_{ij}Q$.

Example. The derivative of a vector field $F_i(\mathbf{r})$ is a tensor $T_{ij} = \frac{\partial F_i}{\partial x_j}$, a tensor field. Our decomposition given above has the symmetric traceless piece

$$P_{ij} = \frac{1}{2}\left(\frac{\partial F_i}{\partial x_j} + \frac{\partial F_j}{\partial x_i}\right) - \frac{1}{3}\delta_{ij}\frac{\partial F_k}{\partial x_k} = \frac{1}{2}\left(\frac{\partial F_i}{\partial x_j} + \frac{\partial F_j}{\partial x_i}\right) - \frac{1}{3}\delta_{ij}\nabla\cdot\mathbf{F},$$

an antisymmetric piece $A_{ij} = \varepsilon_{ijk}B_k$, where

$$B_k = \frac{1}{2}\varepsilon_{ijk}\frac{\partial F_i}{\partial x_j} = -\frac{1}{2}(\nabla\times\mathbf{F})_k.$$

and trace

$$Q = \frac{\partial F_k}{\partial x_k} = \nabla\cdot\mathbf{F}.$$

Hence a complete description involves a scalar $\nabla\cdot\mathbf{F}$, a vector $\nabla\times\mathbf{F}$, and a symmetric traceless tensor P_{ij}.

14.2 The inertia tensor

Consider masses m_α with positions $\mathbf{r}_\alpha$, all rotating with angular velocity $\boldsymbol{\omega}$ about $\mathbf{0}$. So the velocities are $\mathbf{v}_\alpha = \boldsymbol{\omega}\times\mathbf{r}_\alpha$. The total angular momentum is

$$\begin{aligned}\mathbf{L} &= \sum_\alpha \mathbf{r}_\alpha \times m_\alpha \mathbf{v}_\alpha \\ &= \sum_\alpha m_\alpha \mathbf{r}_\alpha \times (\boldsymbol{\omega}\times\mathbf{r}_\alpha) \\ &= \sum_\alpha m_\alpha(|\mathbf{r}_\alpha|^2\boldsymbol{\omega} - (\mathbf{r}_\alpha\cdot\boldsymbol{\omega})\mathbf{r}_\alpha).\end{aligned}$$

by vector identities. In components, we have

$$L_i = I_{ij}\omega_j,$$

where

Definition (Inertia tensor)**.** The *inertia tensor* is

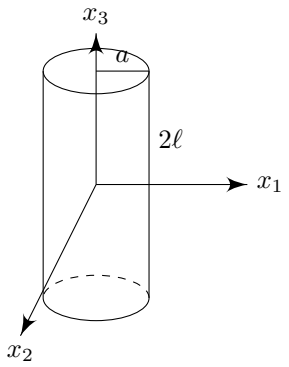
$$I_{ij} = \sum_\alpha m_\alpha[|\mathbf{r}_\alpha|^2\delta_{ij} - (\mathbf{r}_\alpha)_i(\mathbf{r}_\alpha)_j].$$

For a rigid body occupying volume V with mass density $\rho(\mathbf{r})$, we replace the sum with an integral to obtain

$$I_{ij} = \int_V \rho(\mathbf{r})(x_k x_k\delta_{ij} - x_i x_j)\,\mathrm{d}V.$$

By inspection, I is a symmetric tensor.

Example. Consider a rotating cylinder with uniform density ρ_0. The total mass is $2\ell\pi a^2\rho_0$.



Use cylindrical polar coordinates:

$$\begin{aligned} x_1 &= r\cos\theta \\ x_2 &= r\sin\theta \\ x_3 &= x_3 \\ \mathrm{d}V &= r\ \mathrm{d}r\ \mathrm{d}\theta\ \mathrm{d}x_3 \end{aligned}$$

We have

$$\begin{aligned} I_{33} &= \int_V \rho_0(x_1^2 + x_2^2)\ \mathrm{d}V \\ &= \rho_0 \int_0^a \int_0^{2\pi} \int_{-\ell}^{\ell} r^2(r\ \mathrm{d}r\ \mathrm{d}\theta\ \mathrm{d}x_2) \\ &= \rho_0 \cdot 2\pi \cdot 2\ell \left[\frac{r^4}{4}\right]_0^a \\ &= \varepsilon_0 \pi \ell a^4. \end{aligned}$$

Similarly, we have

$$\begin{aligned} I_{11} &= \int_V \rho_0(x_2^2 + x_3^2)\ \mathrm{d}V \\ &= \rho_0 \int_0^a \int_0^{2\pi} \int_{-\ell}^{\ell} (r^2\sin^2\theta + x_3^2) r\ \mathrm{d}r\ \mathrm{d}\theta\ \mathrm{d}x_3 \\ &= \rho_0 \int_0^a \int_0^{2\pi} r\left(r^2\sin^2\theta\,[x_3]_{-\ell}^{\ell} + \left[\frac{x_3^3}{3}\right]_{-\ell}^{\ell}\right)\ \mathrm{d}\theta\ \mathrm{d}r \\ &= \rho_0 \int_0^a \int_0^{2\pi} r\left(r^2\sin^2\theta 2\ell + \frac{2}{3}\ell^3\right)\ \mathrm{d}\theta\ \mathrm{d}r \\ &= \rho_0 \left(2\pi a \cdot \frac{2}{3}\ell^3 + 2\ell \int_0^a r^2\ \mathrm{d}r \int_0^{2\pi} \sin^2\theta\right) \\ &= \rho_0 \pi a^2 \ell \left(\frac{a^2}{2} + \frac{2}{3}\ell^2\right) \end{aligned}$$

By symmetry, the result for I_{22} is the same.

How about the off-diagonal elements?

$$\begin{aligned} I_{13} &= -\int_V \rho_0 x_1 x_3 \, \mathrm{d}V \\ &= -\rho_0 \int_0^a \int_{-\ell}^{\ell} \int_0^{2\pi} r^2 \cos\theta x_3 \, \mathrm{d}r \, \mathrm{d}x_3 \, \mathrm{d}\theta \\ &= 0 \end{aligned}$$

Since $\int_0^{2\pi} \mathrm{d}\theta \cos\theta = 0$. Similarly, the other off-diagonal elements are all 0. So the non-zero components are

$$\begin{aligned} I_{33} &= \frac{1}{2} M a^2 \\ I_{11} = I_{22} &= M\left(\frac{a^2}{4} + \frac{\ell^2}{3}\right) \end{aligned}$$

In the particular case where $\ell = \frac{a\sqrt{3}}{2}$, we have $I_{ij} = \frac{1}{2} m a^2 \delta_{ij}$. So in this case,

$$\mathbf{L} = \frac{1}{2} M a^2 \boldsymbol{\omega}$$

for rotation about any axis.

14.3 Diagonalization of a symmetric second rank tensor

Recall that using matrix notation,

$$T = (T_{ij}), \quad T' = (T'_{ij}), \quad R = (R_{ij}),$$

and the tensor transformation rule $T'_{ij} = R_{ip} R_{jq} T_{pq}$ becomes

$$T' = RTR^T = RTR^{-1}.$$

If T is symmetric, it can be diagonalized by such an orthogonal transformation. This means that there exists a basis of orthonormal eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ for T with real eigenvalues $\lambda_1, \lambda_2, \lambda_3$ respectively. The directions defined by $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are the *principal axes* for T, and the tensor is diagonal in Cartesian coordinates along these axes.

This applies to any symmetric rank-2 tensor. For the special case of the inertia tensor, the eigenvalues are called the *principal moments of inertia*.

As exemplified in the previous example, we can often guess the correct principal axes for I_{ij} based on the symmetries of the body. With the axes we chose, I_{ij} was found to be diagonal by direct calculation.

15 Invariant and isotropic tensors

15.1 Definitions and classification results

Definition (Invariant and isotropic tensor)**.** A tensor T is *invariant* under a particular rotation R if

$$T'_{ij\cdots k} = R_{ip}R_{jq}\cdots R_{kr}T_{pq\cdots r} = T_{ij\cdots k},$$

i.e. every component is unchanged under the rotation.

A tensor T which is invariant under every rotation is *isotropic*, i.e. the same in every direction.

Example. The inertia tensor of a sphere is isotropic by symmetry.

δ_{ij} and ε_{ijk} are also isotropic tensors. This ensures that the component definitions of the scalar and vector products $\mathbf{a}\cdot\mathbf{b} = a_ib_j\delta_{ij}$ and $(\mathbf{a}\times\mathbf{b})_i = \varepsilon_{ijk}a_jb_k$ are independent of the Cartesian coordinate system.

Isotropic tensors in $\mathbb{R}^3$ can be classified:

Theorem.

(i) There are no isotropic tensors of rank 1, except the zero tensor.

(ii) The most general rank 2 isotropic tensor is $T_{ij} = \alpha\delta_{ij}$ for some scalar α.

(iii) The most general rank 3 isotropic tensor is $T_{ijk} = \beta\varepsilon_{ijk}$ for some scalar β.

(iv) All isotropic tensors of higher rank are obtained by combining δ_{ij} and ε_{ijk} using tensor products, contractions, and linear combinations.

We will provide a sketch of the proof:

Proof. We analyze conditions for invariance under specific rotations through π or $\pi/2$ about coordinate axes.

(i) Suppose T_i is rank-1 isotropic. Consider a rotation about x_3 through π:

$$(R_{ij}) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We want $T_1 = R_{ip}T_p = R_{11}T_1 = -T_1$. So $T_1 = 0$. Similarly, $T_2 = 0$. By consider a rotation about, say x_1, we have $T_3 = 0$.

(ii) Suppose T_{ij} is rank-2 isotropic. Consider

$$(R_{ij}) = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which is a rotation through $\pi/2$ about the x_3 axis. Then

$$T_{13} = R_{1p}R_{3q}T_{pq} = R_{12}R_{33}T_{23} = T_{23}$$

and

$$T_{23} = R_{2p}R_{3q}T_{pq} = R_{21}R_{33}T_{13} = -T_{13}$$

So $T_{13} = T_{23} = 0$. Similarly, we have $T_{31} = T_{32} = 0$.

We also have

$$T_{11} = R_{1p}R_{1q}T_{pq} = R_{12}R_{12}T_{22} = T_{22}.$$

So $T_{11} = T_{22}$.

By picking a rotation about a different axis, we have $T_{21} = T_{12}$ and $T_{22} = T_{33}$.

Hence $T_{ij} = \alpha\delta_{ij}$.

(iii) Suppose that T_{ijk} is rank-3 isotropic. Using the rotation by π about the x_3 axis, we have

$$T_{133} = R_{1p}R_{3q}R_{3r}T_{pqr} = -T_{133}.$$

So $T_{133} = 0$. We also have

$$T_{111} = R_{1p}R_{1q}R_{1r}T_{pqr} = -T_{111}.$$

So $T_{111} = 0$. We have similar results for π rotations about other axes and other choices of indices.

Then we can show that $T_{ijk} = 0$ unless all i, j, k are distinct.

Now consider

$$(R_{ij}) = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

a rotation about x_3 through $\pi/2$. Then

$$T_{123} = R_{1p}R_{2q}R_{3r}T_{pqr} = R_{12}R_{21}R_{33}T_{213} = -T_{213}.$$

So $T_{123} = -T_{213}$. Along with similar results for other indices and axes of rotation, we find that T_{ijk} is totally antisymmetric, and $T_{ijk} = \beta\varepsilon_{ijk}$ for some β. □

Example. The most general isotropic tensor of rank 4 is

$$T_{ijk\ell} = \alpha\delta_{ij}\delta_{k\ell} + \beta\delta_{ik}\delta_{j\ell} + \gamma\delta_{i\ell}\delta_{jk}$$

for some scalars α, β, γ. There are no other independent combinations. (we might think we can write a rank-4 isotropic tensor in terms of ε_{ijk}, like $\varepsilon_{ijp}\varepsilon_{k\ell p}$, but this is just $\delta_{ik}\delta_{j\ell} - \delta_{i\ell}\delta_{jk}$. It turns out that anything you write with ε_{ijk} can be written in terms of δ_{ij} instead)

15.2 Application to invariant integrals

We have the following very useful theorem. It might seem a bit odd and arbitrary at first sight — if so, read the example below first (after reading the statement of the theorem), and things will make sense!

Theorem. Let

$$T_{ij\cdots k} = \int_V f(\mathbf{x}) x_i x_j \cdots x_k \, \mathrm{d}V.$$

where $f(\mathbf{x})$ is a scalar function and V is some volume.

Given a rotation R_{ij}, consider an *active* transformation: $\mathbf{x} = x_i \mathbf{e}_i$ is mapped to $\mathbf{x}' = x_i' \mathbf{e}_i$ with $x_i' = R_{ij} x_i$, i.e. we map the components but not the basis, and V is mapped to V'.

Suppose that under this active transformation,

(i) $f(\mathbf{x}) = f(\mathbf{x}')$,

(ii) $V' = V$ (e.g. if V is all of space or a sphere).

Then $T_{ij\cdots k}$ is invariant under the rotation.

Proof. First note that the Jacobian of the transformation R is 1, since it is simply the determinant of R ($x_i' = R_{ip} x_p \Rightarrow \frac{\partial x_i'}{\partial x_p} = R_{ip}$), which is by definition 1. So $\mathrm{d}V = \mathrm{d}V'$.

Then we have

$$\begin{aligned}
R_{ip} R_{jq} \cdots R_{kr} T_{pq\cdots r} &= \int_V f(\mathbf{x}) x_i' x_j' \cdots x_k' \, \mathrm{d}V \\
&= \int_V f(\mathbf{x}') x_i' x_j' \cdots x_k' \, \mathrm{d}V \qquad \text{using (i)} \\
&= \int_{V'} f(\mathbf{x}') x_i' x_j' \cdots x_k' \, \mathrm{d}V' \qquad \text{using (ii)} \\
&= \int_V f(\mathbf{x}) x_i x_j \cdots x_k \, \mathrm{d}V \qquad \text{since } x_i \text{ and } x_i' \text{ are dummy} \\
&= T_{ij\cdots k}. \qquad \square
\end{aligned}$$

The result is particularly useful if (i) and (ii) hold for *any* rotation R, in which case $T_{ij\cdots k}$ is isotropic.

Example. Let

$$T_{ij} = \int_V x_i x_j \, \mathrm{d}V,$$

with V being a solid sphere of $|\mathbf{r}| < a$. Our result applies with $f = 1$, which, being a constant, is clearly invariant under rotations. Also the solid sphere is invariant under any rotation. So T must be isotropic. But the only rank 2 isotropic tensor is $\alpha \delta_{ij}$. Hence we must have

$$T_{ij} = \alpha \delta_{ij},$$

and all we have to do is to determine the scalar α.

Taking the trace, we have

$$T_{ii} = 3\alpha = \int_V x_i x_i \, \mathrm{d}V = 4\pi \int_0^a r^2 \cdot r^2 \, \mathrm{d}r = \frac{4}{5} \pi a^5.$$

So

$$T_{ij} = \frac{4}{15} \pi a^5 \delta_{ij}.$$

Normally if we are only interested in the $i \neq j$ case, we just claim that $T_{ij} = 0$ by saying "by symmetry, it is 0". But now we can do it (more) rigorously!

There is a closely related result for the inertia tensor of a solid sphere of constant density ρ_0, or of mass $M = \frac{4}{3}\pi a^3 \rho_0$.

Recall that

$$I_{ij} = \int_V \rho_0 (x_k x_k \delta_{ij} - x_i x_j) \, \mathrm{d}V.$$

We see that I_{ij} is isotropic (since we have just shown that $\int x_i x_j \, \mathrm{d}V$ is isotropic, and $x_k x_k \delta_{ij}$ is also isotropic). Let $I_{ij} = \beta \delta_{ij}$. Then

$$\begin{aligned}
I_{ij} &= \int_V \rho_0 (x_k x_k \delta_{ij} - x_i x_j) \, \mathrm{d}V \\
&= \rho_0 \left(\delta_{ij} \int_V x_k x_k \, \mathrm{d}V - \int_V x_i x_j \, \mathrm{d}V \right) \\
&= \rho_0 \left(\delta_{ij} T_{kk} - T_{ij} \right) \\
&= \rho_0 \left(\frac{4}{5} \pi a^5 \delta_{ij} - \frac{4}{15} \pi a^5 \delta_{ij} \right) \\
&= \frac{8}{15} \rho_0 \pi a^5 \delta_{ij} \\
&= \frac{2}{5} M a^2 \delta_{ij}.
\end{aligned}$$